Abstract

The investigations in this thesis are motivated by the discovery of a quasi-biennial fluctuation of the sunspot umbra-penumbra area ratio (U/P) in the Debrecen Photoheliographic Data (DPD), a detailed catalogue of sunspot activity. A summary of the phenomenon and the analyses of its periodic behaviour are followed by a presentation of some ideas of possible physical interpretations. The emphasis of the thesis lies in a wide survey of solar mid-term periodicities and their possible connections to the U/P fluctuation. When relevant data have been accessible, the periodic behaviours have been explored by filtering methods and spectral analyses. Several separate quantitative comparisons to the U/P fluctuation are given, illustrated by composite figures.

No firm conclusion about the origin of the U/P fluctuation can be drawn at present, but several interesting similarities of the phenomenon and other mid-term periodicities are revealed.

The research has been carried out at the Heliophysical Observatory in Debrecen, Hungary.

Preface

This preface does not only appear in my thesis but also in the thesis written by Csilla Szasz, called *Fluctuations of the Sunspot Umbra-Penumbra Ratio*. We have composed the preface together since we, both being students at the Master of Science Programme in Space Engineering at Luleå University of Technology (LTU) in Sweden, have accomplished our diploma work in close collaboration. Our research has been carried out at the Heliophysical Observatory of the Hungarian Academy of Sciences, in the city of Debrecen, Hungary, where we have spent five months during the autumn of 2002. Already during the first months of our analyses, we discovered something interesting enough about sunspots to deserve full-time concern from both of us during the rest of our stay. To satisfy the demand of two separate reports from our home institution at LTU, our investigations are divided into two self-reliant parts, compiled to be digestible on their own. For a complete overview, the interested reader is advised to read both theses.

We would first of all like to thank our supervisor András Ludmány at the Heliophysical Observatory for his engagement, help and encouragement during the course of our work. Also Tünde Baranyi has been a great help, taking her time to look at new results and trying to interpret them. Without the help of György Mező, with his experience and patience opening doors to methods and software of which we did not have previous knowledge, the computations would have been much slower. All three of them have helped us with valuable ideas and discussions making our work interesting and fun. We would also like to express our gratitude to László Tóth for taking active part in making our stay at the Heliophysical Observatory unforgettable.

Sverker Fredriksson and Björn Graneli have been our allies at LTU, encouraging and assisting us with the arrangements before we left for Hungary and always eager to help and answer questions during our stay abroad.

Our research has been financed by long-term scholarships from the Hungarian Ministry of Education and support from Ångpanneföreningen's Research Foundation. The journey to Hungary has been funded by travel scholarships from the Seth M. Kempe Foundation.

The content of this thesis has been partly presented at two separate occasions in Hungary. One was a talk held at the 30^{th} national ionospheric and magnetospheric conference in Tihany, November 11-13, 2002 and the other a one-hour seminar at the Astronomical Department of the Eötvös Loránd University in Budapest, November 20, 2002

Contents

1 Introduction

Our view of the Sun is undergoing an inexhaustible process of development. When we think that we know what the Sun is up to, new fabulous features are found, catching our attention. From the new perspective our approaches can be refined and yet another advancement towards understanding our star can follow. The past decade with the launch of important solar observing satellites as Ulyssus, Yohkoh and SOHO has brought great discoveries as well as verifications of theoretical assumptions. In the same rate as questions are answered, new questions arise, calling for additional and more advanced investigations. It is, however, important not to forget the already assembled treasures of information and assure their continuation when raising new funds for development of state-of-the-art technology.

During the course of the diploma work of Csilla Szasz and myself at the Heliophysical Observatory in Debrecen, Hungary, an unexpected quasi-biennial fluctuation of the sunspot umbra-penumbra (U/P) area ratio was found in the Debrecen Photoheliographic Data. The complete description of our investigations leading to this discovery and the data used are presented in the thesis [Sz03], whereas this thesis contains some ideas of possible physical interpretations and a wide survey of solar mid-term periodicities.

As the phenomenon found is a novelty, and the details of sunspot appearance are not yet fully understood, no definite suggestion of an underlying mechanism can be presented. Some hints of how things may come about are given in Section 4 and a general overview of solar periodic behaviour, containing direct comparisons with the U/P ratio when possible, is summarized in Section 5. In the latter section such periodicities as fluctuations of the solar radius, the temperature of the photosphere, the relative sunspot numbers and solar irradiance are discussed. The term 'Quasi-Biennial Oscillation' (QBO), today loosely used in several contexts, has its origin in the field of meteorology and is described in Section 6.

A simple model of the processes through which sunspots are created by the rising of magnetic flux tubes in the convective zone and their appearance in the photosphere is treated in Section 2. For the sake of completeness, a summary of the U/P fluctuation as well as the complete analyses of its periodic behaviour are presented in Section 3.

It is said that nothing is new under the Sun, but this statement is only true for the one who has already looked everywhere.

2 Sunspots as Described by Magnetic Flux Tubes

The knowledge of the structure of sunspots is in a sense not more certain than the assumptions one can make of the structure of a tree by looking at the tree top from above. The leaves and branches hide the underlying stem from direct inspection. Several theoretical attempts have been made to describe the sunspot phenomenon, but it is still not possible to distinguish right from wrong. Different models are used in different contexts.

2.1 Descriptive Introduction to Sunspots

Regions on the Sun with high magnetic activity are called active regions, and it is within these that sunspots most often arise. Sunspots observed in white-light consist typically of a darker, inner region called umbra surrounded by a lighter, outer penumbra. The reason for the sunspots being darker than the surrounding photosphere is that they are cooler. There is a wide variety of sunspot structures. Sunspots are irregular in shape and only very seldom circular, if ever, sometimes an umbra even lacks a penumbra or vice versa. Umbra diameter sizes range from about 2 000 km to far above 20 000 km.

Less dark features without a penumbra and smaller than about 2500 km in diameter are termed pores. These have much shorter lifetimes than spots, about a few tens of minutes. The sunspots live from a couple of hours to weeks.

A granulation structure similar to the photospheric one, but in smaller scale, can be observed within the umbra area. The penumbra consists of penumbral filaments, which, in turn, consist of dark cores inside lighter outer shells.

Sunspots occur most frequently in groups, which can have different characteristics. The Mt. Wilson magnetic classification of sunspots, or rather sunspot groups, was introduced in 1919 and is summarized in Table 1. [Fo90] [Pr84] [Zi88]

Sunspot Class	Characteristics
α	A single, dominant spot
β	A pair of dominant spots of opposite polarity
γ	Complex groups with irregular distributions of polarities
$\beta\gamma$	Bipolar groups with no marked north-south inversion line
δ	Umbrae with opposite polarity in a single penumbra

Table 1: Table of the Mt. Wilson magnetic classification of sunspots. [Zi88]

2.2 Simplified Description of Sunspot Formation

Deep in the convective zone the kinetic energy density of the plasma convective motions is much higher than the magnetic energy density,

$$\frac{\rho \cdot U^2}{2} \gg \frac{B^2}{2\mu}, \tag{1}$$

where ρ is the plasma mass density, U is the characteristic speed, B the magnetic field strength and μ the magnetic permeability. Furthermore the magnetic Reynold's number R_m can be much larger than unity,

$$R_m = \frac{U \cdot L}{\eta} \gg 1 \tag{2}$$

where L is the characteristic length scale and η is the magnetic diffusivity. In these cases the magnetic field lines behave as if they were frozen into the plasma. They move with it and can be stretched and wound up by the plasma motions. These processes continue until the magnetic field lines are stretched so much that the magnetic energy density exceeds the kinetic energy density or the local magnetic Reynold's number decreases to an order of unity.

Many magnetic phenomena of the sun, such as the overall structure of sunspots and prominences, can be described by magnetohydrostatic equations since they appear stationary on time-scales comparable to the Alfvén travel time τ_A given by

$$\tau_A = L \cdot \frac{1}{v_A} = L \cdot \frac{\sqrt{\mu \cdot \rho_0}}{B_0}, \tag{3}$$

where L once again is the characteristic length scale of the phenomenon, v_A the Alfvén velocity and ρ_0, B_0 are typical values of the plasma density and the magnetic field strength. The characteristics of a magnetic flux tube in hydrostatic pressure equilibrium with its surroundings are governed by

$$\frac{k_B \cdot T_e \cdot \rho_e}{m_e} = \frac{k_B \cdot T_i \cdot \rho_i}{m_i} + \frac{B_i^2}{2\mu}, \tag{4}$$

where k_B is the Boltzmann's constant, T the temperature, ρ the density, m the mean particle mass and B the magnetic field strength, all with indices e and i corresponding to external and internal properties of the flux tube. The magnetic field strength outside flux tubes is low enough for the external magnetic pressure to be neglected in comparison with the dominating plasma gas pressure, appearing alone on the left-hand side of Equation 4. This is not the case inside flux tubes where the contributions from the magnetic pressure, given by the second term on the right-hand side of Equation 4, are significant. The flux tube will hereby due to equilibrium conditions have an internal density and temperature lower than its surroundings since the external gas pressure on the left-hand side of Equation 4 has to balance the sum of the internal gas pressure and the internal magnetic pressure on the right-hand side.

If the flux tube has lower density than its surroundings it will start rising due to gravitational forces, as illustrated in Figure 1.

The rising process of flux tubes can continue as long as the buoyancy forces on the left-hand side of

$$(\rho_e - \rho_i) \cdot g_{loc} > \frac{B_i^2}{\mu \cdot l} \tag{5}$$

are larger than the restoring magnetic tension on the right-hand side, caused by the curvature of the field lines,where g_{loc} is the local gravitational constant and l is the distance between the footprints of the flux tube in the convective zone. The magnetic field strength of the flux tube decreases as the tube reaches the photospheric surface where the with height decreasing external temperature allows the flux tube, still considered to be in horizontal hydrostatic equilibrium, to fan out. Depending on the distance between the footprints of the flux tube, the vertically rising tube can decelerate to an equilibrium where the magnetic buoyancy equals the magnetic tension, since the latter is a function of the radius of curvature. Where the flux tube penetrates the visible photospheric surface, sunspots are seen having the characteristics of the flux tube: a temperature deficit and an excess of magnetic field strength compared to the surrounding plasma.

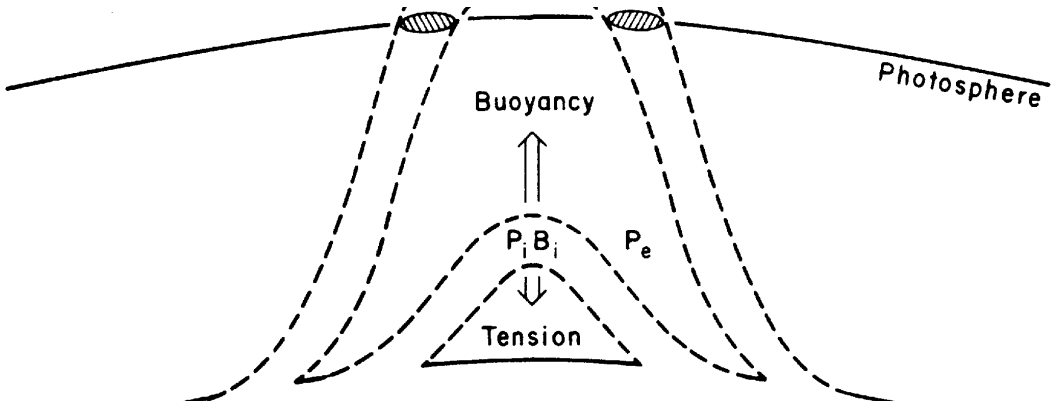

Figure 1: Schematic sketch of an emerging flux tube [Pr84]. Where the Ω-shaped flux tube penetrates the photospheric surface, sunspots are seen having the characteristic temperature deficiency and excess of magnetic field strength compared to the surrounding photosphere.

2.3 Monolithic and Cluster Models

There are basically three different theoretical models for how sunspots are formed by flux tubes and maintained stable in the photosphere. The classical picture is that a sunspot is formed by a single magnetic flux tube, which makes out the base of the magnetohydrostatic and magnetoconvective models. These models are called monolithic and one expects that the vertical dimensions of a sunspot is of the same order as the horizontal dimensions, as illustrated in Figure 2, from Parker [Pa79]. The third theoretical model is called the cluster model.

The monolithic models of sunspots are developed on a well-defined set of mathematical equations, but have difficulties in explaining why the surface temperature of ∼3900 K and the magnetic field strength of 3 kG seem to be independent of sunspot dimensions. The cluster model introduced by Parker in 1979 [Pa79] is on the other hand not firmly developed on a mathematical basis, but explains the qualitative features mentioned well.

In the cluster model, a sunspot is made up of a bunch of flux tubes held together by buoyancy forces due to the fanning out of field lines in the upper photosphere, appearing as one single visible unit. A qualitative sketch by Parker is presented in Figure 3. Any small flux tube that tries to separate from the unified cluster in the upper photosphere has to move outward and downward against the stabilizing buoyancy forces. Further down, where the sides of the flux tubes decline to the vertical direction the buoyancy forces can no longer stabilize the cluster, and the flux tubes break apart. The depth where this happens is considered approximately the same for all sizes of sunspots. There might also be a characteristic size of the separate flux tubes involved, which might explain the independence of sunspot size to parameters as surface temperature and magnetic field strength.

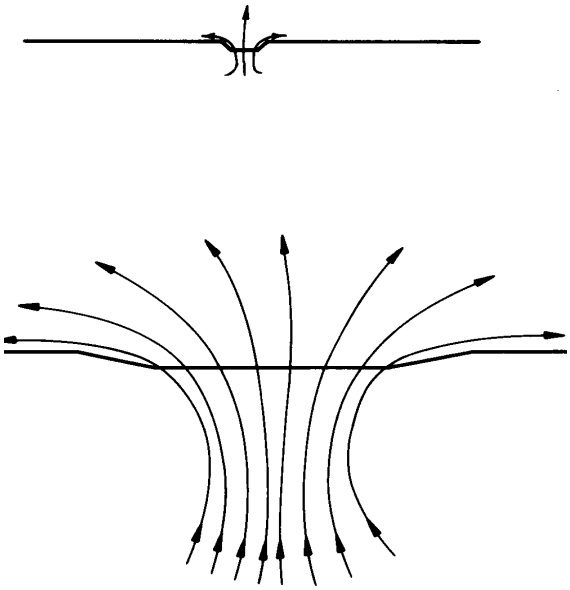

Figure 2: Cross-section sketch illustrating the scaling of the magnetic structure of sunspots according to monolithic sunspot models [Pa79].

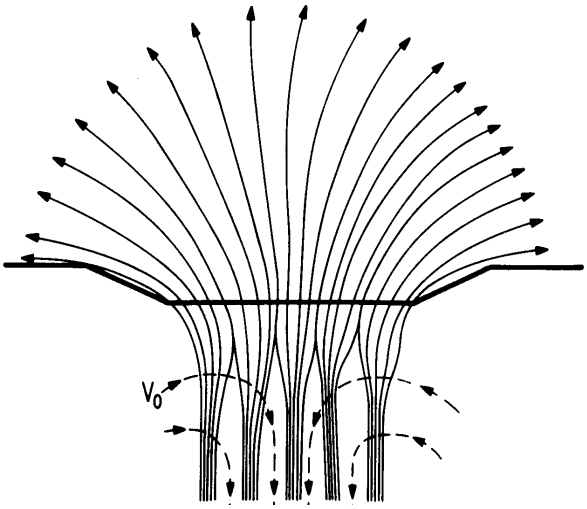

Figure 3: Cross-section of the magnetic structure of a large sunspot according to the cluster model [Pa79].

3 The Umbra-Penumbra (U/P) Area Ratio

3.1 Debrecen Sunspot Catalogues

The Debrecen Photoheliographic Data (DPD) is a catalogue containing sunspot parameters. The production of the DPD is based on analyses of full-disk solar photographic plates in the continuous visible spectrum of light, or white-light for short. Parameters included in the data set are those extractable from full-disk solar photographic plates, as position and area of the umbra and penumbra, not only for sunspot groups but also for individual spots. It is the only sunspot catalogue taking into account every spot observed, even the smallest ones. The catalogue contains observations for each day of the published years. Usually around 300 days a year can be covered by observational material from the Debrecen Observatory itself or its observing station on top of a 43 meter high water tower in the Hungarian city of Gyula. Observational material for the rest of the year has to be searched for and provided by other observatories around the world.

The ancestor of the DPD is the Greenwich Photoheliographic Results, being the classical catalogue containing sunspot characteristics for every day from the year 1874 to 1976. Beginning with 1977, the programme was taken over by the Debrecen Heliophysical Observatory, and the Debrecen Photoheliographic Results (DPR) was born. Being a more complex catalogue than the one it is the continuation of, the process of compiling data is slow. This delay gave rise to the second Debrecen catalogue already mentioned, the DPD. Preparations for the production of the DPD began in 1992 and the first compiled year of the series was introduced in 1996.

3.1.1 Fluctuation of the DPD U/P Area Ratio

The monthly mean umbra-penumbra area ratio (U/P) with all spots included, except spots with zero umbra or penumbra area and with the between spots shared umbra areas added to one spot, is presented in Figure 4. A period of about two years can be seen. The error-bars correspond to one estimated standard error of mean $S_{\bar{x}}$ defined by

$$S_{\bar{x}} = \frac{S}{\sqrt{N}}, \tag{6}$$

where S is the sample standard deviation of the epoch and N is the sample size.

The data for 1986 and first half of 1987 are produced by a different method of measurement than the rest of the data, which can explain the large spread [Sz03]. The spread may likewise partly be caused by the low amount of sunspots visible, as this period of the data set is governed at sunspot minimum. The deviating value in late 1996 contains only three small spots, which means that its location off the curve is uncertain. The swerving value in mid-1995 though, is calculated on 69 spots but since it occurs during a maximum, its deviation is not as significant.

There is one monthly mean missing, May 1986. During this month no sunspots were visible.

3.1.2 Fluctuation of the DPR U/P Area Ratio

The analyses of the monthly mean of the U/P area ratio for sunspot groups in the DPR for the available years of 1977 and 1978 give a one-year period in the manner indicated in Figure 5. There is a discrepancy in the period when comparing the DPR to the DPD,

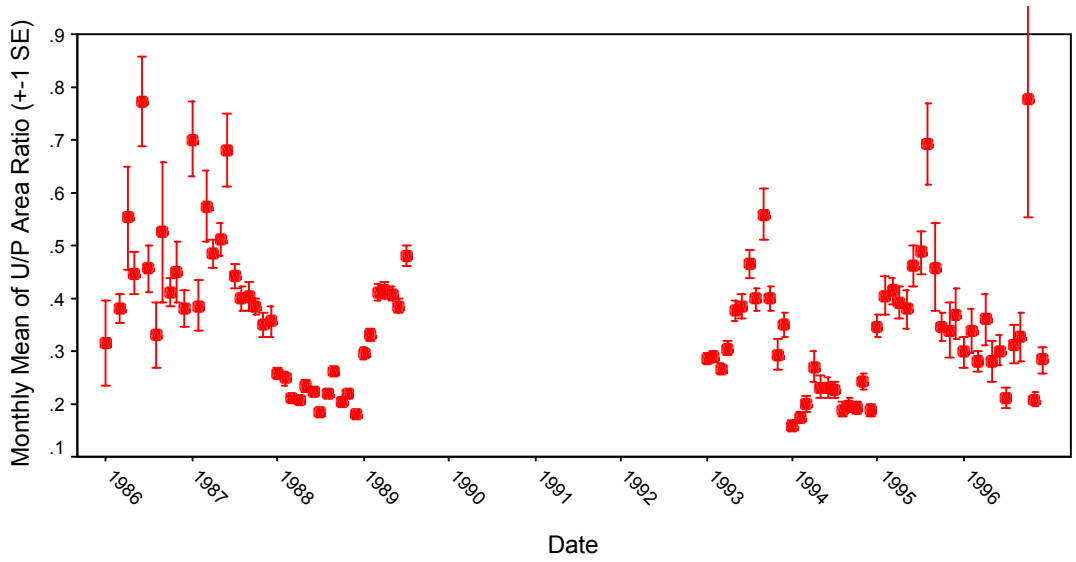

Figure 4: Monthly mean U/P area ratio for sunspots in the DPD. All spots with both a distinguishable umbra and penumbra are included. The error-bars correspond to one standard error.

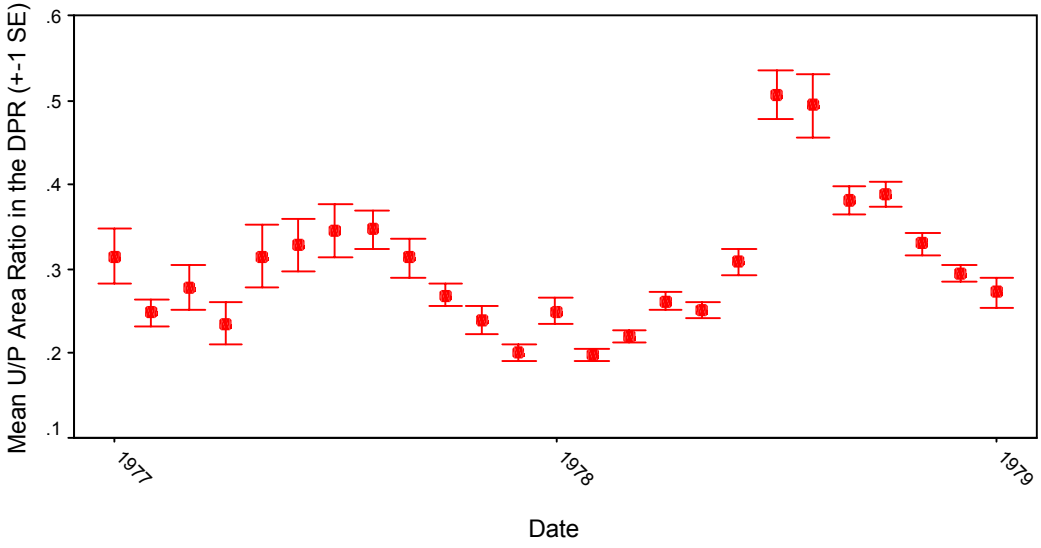

Figure 5: Monthly mean U/P area ratio for sunspot groups in the DPR. The error-bars correspond to one standard error.

but a consistency in amplitude. However, the shorter period in the DPR does not have to contradict the two-year period found in the DPD since the length of the cycle may vary.

3.2 Analyzing the Periodic Behaviour of the DPD U/P Area Ratio

This section will describe the basic features of the U/P fluctuation found in the DPD. We have investigated the U/P ratio as a function of sunspot size as well as performed different methods to analyze its behaviour in time.

3.2.1 The U/P Area Ratio as a Function of Umbra Area

Figure 6 shows the U/P area ratio as a function of umbra area. We have divided the data set roughly into fluctuation minima and maxima, as the state of the cycle can be interpreted from the plot in Figure 4, by separating the minimum years of 1988 and 1994 from the maximum years of 1993 and 1995. In Figure 6 the U/P area ratio is hereafter plotted as a function of umbra area for the different years separately. The periodicity can be seen to be caused by spots with an umbra area smaller than approximately 16 millionths of the solar hemisphere.

Almost all spots have a uniform U/P area ratio during minima, whereas spots with an umbra area smaller than 16 millionths of the solar hemisphere clearly diverge to higher U/P values during maxima. The deviation increases with decreasing size of the umbra. There may be a tendency for small spots to have a departing U/P ratio also during minima, but this deviation is of a much lower order than at maxima. The deviation during minima is small enough to perhaps be caused by measurement uncertainties as areas of sunspots in the catalogue are rounded to integer values. The rounding error hereby introduced is larger for smaller spots.

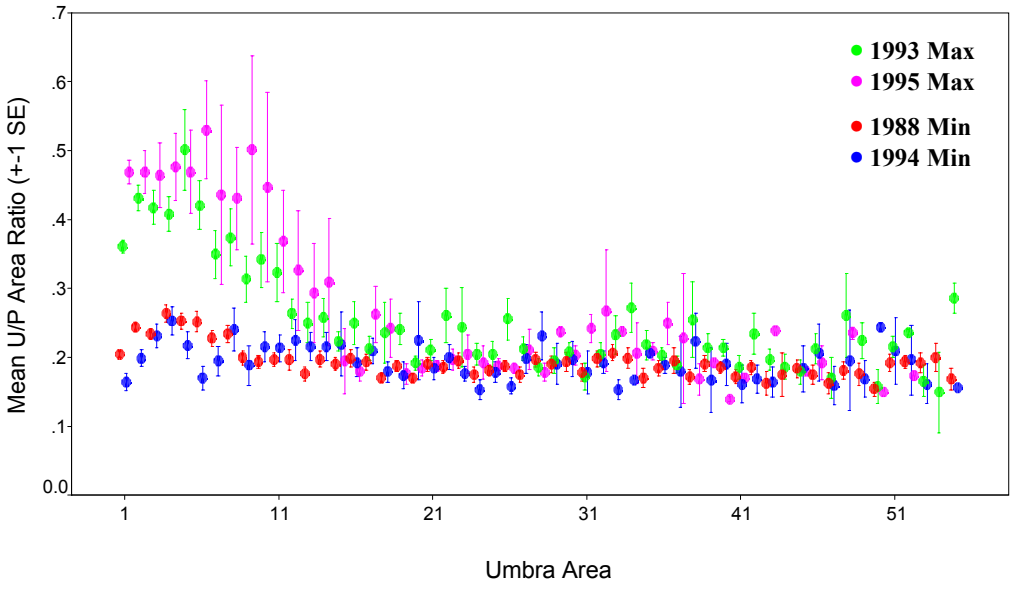

Figure 6: The U/P area ratio as a function of umbra area, which is given in millionths of the solar hemisphere. The data set is divided roughly into fluctuation minima and maxima by separating the years of 1988 and 1994 as well as the years of 1993 and 1995.

3.2.2 Spectral Analysis of the U/P Area Ratio Fluctuation

There are several different methods of investigating the spectral contents of a discrete signal. With the help of Matlab we have analyzed the periodic behaviour of the U/P area ratio fluctuations found in the DPD by estimating its power spectral density (PSD) using periodogram. The two different time intervals, 1986-1989 and 1993-1996, have been treated separately since the Fast Fourier Transform (FFT) requires a sampled signal without gaps to give an accurate result. The samples used are for the sake of simplicity the monthly means presented in Figure 4 and Section 3.1.1. The error introduced by the varying sample interval (due to varying number of days in the months of the year) is small in comparison with the crudity of probing a quasi-biennial signal in a four-year long data set.

The highest analyzable frequency has a period of twice the sampling period according to the Nyqvist sampling theorem. A higher sampling frequency than of the order one month is not suitable for this data set, since the number of gaps due to lack of observed sunspots increases fast with shorter averaging intervals. As we are interested in deducing the characteristics of a quasi-biennial period, the frequency range from zero to six cycles per year is enough. In place of the one and only missing monthly mean of May 1986, the mean of the two neighboring monthly values is used due to the need of a consistent time series.

In Figure 7 the PSD estimate for the time interval 1986 to 1989 can be seen, and in Figure 8 the estimate for 1993 to 1996. The mean U/P area ratios of the intervals have been subtracted from the two sets of monthly values, respectively, to center the data around zero and reveal their periodic behaviour. The spectral peaks are broad due to the shortness of the two data sets, but we are able to make rough estimations for their frequency contents.

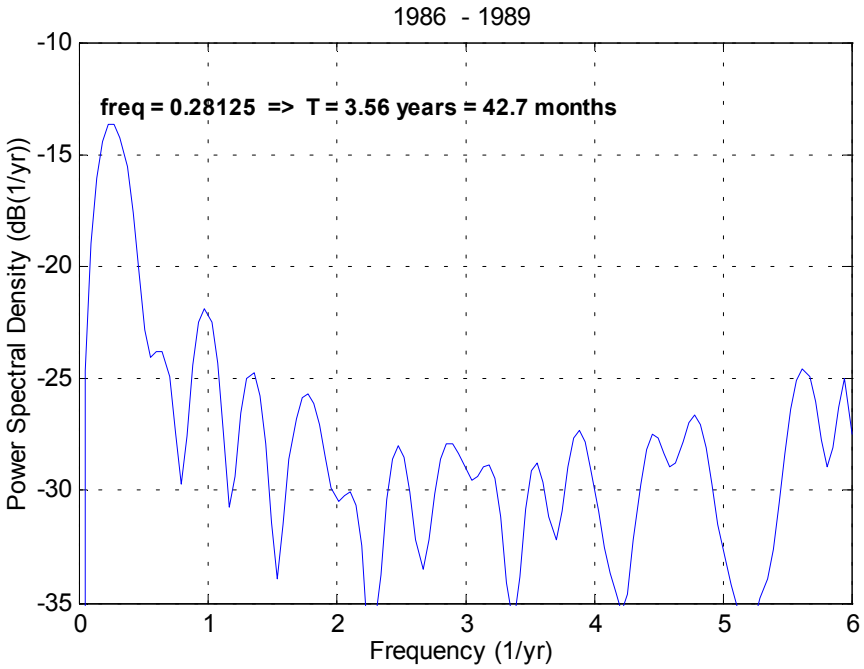

Figure 7: PSD estimate for the monthly mean series of the U/P area ratio for 1986-1989. We have centred the observed fluctuation around zero by subtracting the mean value prior to analysis.

The 1986 to 1989 data set has its highest spectral peak at a frequency corresponding to a period of almost 43 months. This can be visualized by fitting an imagined sinusoidal curve with a 43-month-period to the minimum of 1988 and the high monthly values of 1986 and early 1987 in Figure 4. This curve would have a very high amplitude compared to any more rapidly varying curve fitted to the low values of 1986 and therefore also have higher energy density. The peak is broadened in the sense of shorter periods and has a bump at a frequency corresponding to a period of about 20 months. Since the data of 1986 and first half of 1987 are digitized with the use of a different method and contains a much larger spread of the monthly averages than the following data, we find the period of 43 months untrustworthy. The broadening of the peak towards shorter periods and the bump are signs of a quasi-biennial signal with shorter period than 43 months.

The 1993-1996 spectral analysis in Figure 8 has its highest spectral peak at a frequency corresponding to approximately 21 months. This peak has a lower absolute value than the highest peak of Figure 7 and is broadened in the sense of longer periods. There is a local minimum limiting the peak to only contain power from periods longer than one and a half years.

In both Figure 7 and 8 there are significant peaks also at a frequency corresponding to exactly one year. This might be a sign of observational seasonal effects since the period of one Earth-rotation around the Sun is not expected to appear in a feature of solar origin. The Earth's contribution to the angular momentum of our solar system is negligible and we are not aware of any other ways a planet can influence the Sun such that a solar periodicity with a period equal to the planet's period of revolution is produced.

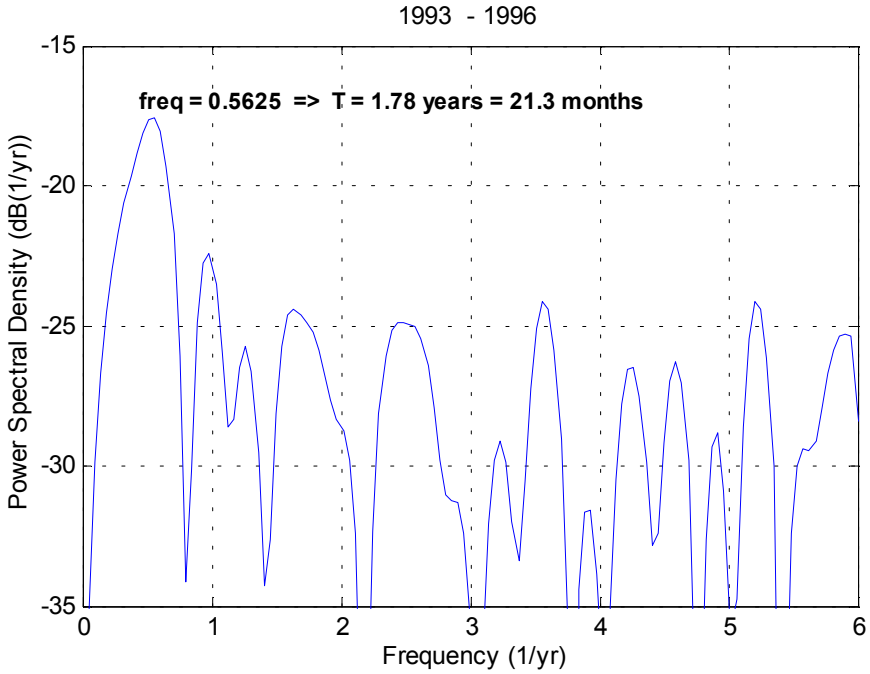

Figure 8: PSD estimate for the monthly mean series of the umbra-penumbra area ratio for 1993 to 1996. We have centred the observed fluctuation around zero by subtracting the mean value before the analysis.

3.2.3 A Bridge Over the Gap?

If we, in spite of the different periods found by spectral analysis in Section 3.2.2, consider the fluctuation as having a more or less constant length of its cycle as it progresses in time, we should be able to fit the two data sets with a single sinusoidal curve. This curve is a prediction of what the U/P area ratio can be expected to be in the gap between the two data sets. We have made a non-linear regression analysis with the statistical software package SPSS, by assuming a model with four free parameters according to the formula

$$U/P = C + A \cdot \sin\left(\omega \cdot time + \phi\right), \tag{7}$$

where the U/P area ratio versus time (in years) is modelled by an arbitrary constant C added to a sinusoidal function with arbitrary amplitude A, angular frequency ω and phase ϕ.

Parameter	**Value**	
Constant (C)	0.3157	
Amplitude (A)	0.1106	
Frequency (ω)	3.117	yr^{-1}
Phase (ϕ)	1.80	rad
Period ($\frac{2\pi}{w} \cdot 12$)	24.2	months

Table 2: Parameter values for the best sinusoidal fit of the monthly mean U/P area ratio of sunspots in the DPD.

The parameter values found in Table 2 are presented with four significant digits, with exception of the phase that is presented with three. The digits are significant in the sense that they are non-changing under different, reasonable starting assumptions. For enabling use of the Levenberg-Marquardt algorithm implemented in SPSS no constraints can be put on the parameters. Without constraints the phase can assume any value with a multiple of 2π added to its fundamental value between 0 and 2π, still giving an identical solution. Thus, different starting conditions give different values for the phase. The value for the phase presented in Table 2 is found by subtracting any multiples of 2π and correcting with π if the amplitude A is negative. (The multiple can contain an odd number of π for an identical solution, but in this case the corresponding amplitude found by the algorithm is negative.)

The model described in Table 2 is illustrated in Figure 9. It can be seen that the sinusoidal with a period of 24.2 months quite well fits the data with exception of 1986 and first half of 1987. This might be due to the non-fitting data not being reliable, the cycle length of the periodicity varying, or, most likely, a combination of the two features.

3.2.4 Variation of the Cycle Length?

There are many different solar mid-term periodicities found during the last decades, as will be reviewed in Section 5. Some of these are less well-established than others, but even the most well-known seem to have an evolving cycle length. The main 11-year sunspot cycle has also been varying between less than ten to an almost twelve years length during the last 140 years [La00], for which reliable observations have been made.

The behaviour of many discovered phenomena of our star are as dynamic as the way they reveal themselves, like the gigantic outbursts of coronal mass ejections. There is

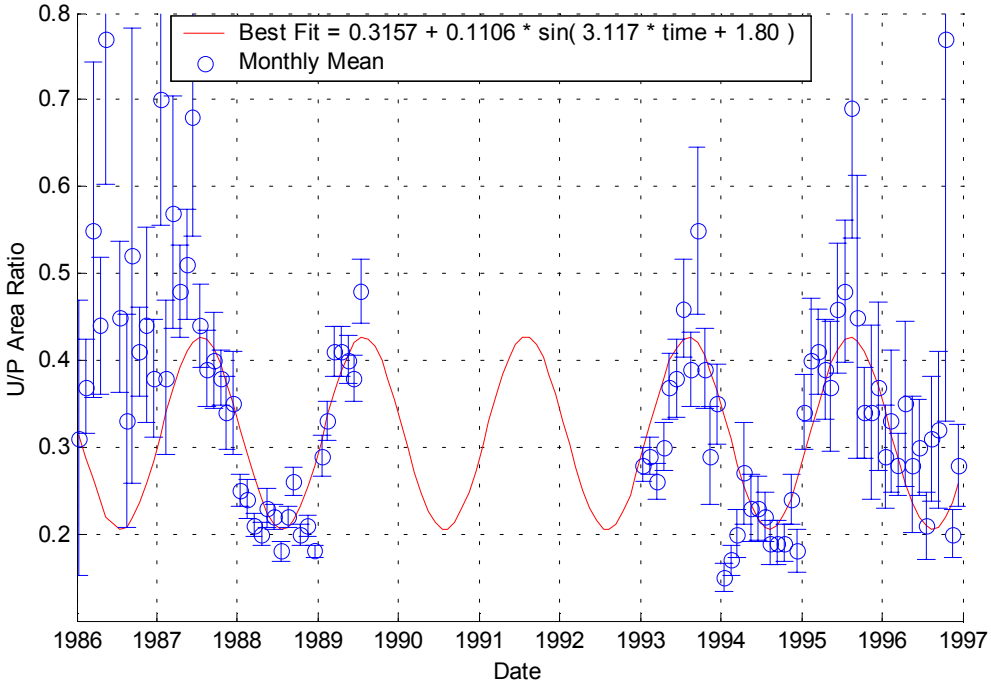

Figure 9: Monthly mean U/P area ratio for sunspots and the best sinusoidal fit found by nonlinear regression with SPSS. The error-bars correspond to two standard errors.

no reason to expect a complete regularity in the case of an umbra-penumbra fluctuation, even though the appearance of sunspots in the photosphere observed in white-light is a quite quiescent phenomenon.

To be able to reveal the main trends of the cyclic behaviour of the U/P area ratio, we have treated the two data sets separately and fitted each of them with least-square-methods in Matlab. The data series are first normalized according to

$$\hat{X} = \frac{U/P - \mu_1}{\mu_2}, \tag{8}$$

where $\hat{X}$ is the normalized data, U/P the original data, μ_1 the mean value and μ_2 the standard deviation. Then a least-square fit is made whereafter the fitted curve is transformed back with the inversion of Equation 8 to enable comparisons with the original data. We have by inspection found a polynomial degree of six to be suitable for the curves to show the quasi-biennial trend as clearly as possible, but not contain faster variations similar to the scatter of the original data.

Plots for both time intervals are shown in Figure 10. The data from 1986 and the first half of 1987 may be uncertain, as discussed previously. The monthly mean of October 1996 has not been taken into account when fitting polynomials to the time interval 1993 to 1996 since it has large error-bars but due to its deviant value strongly influences the polynomial fit.

It can be seen in Figure 10 that the cycle length may be shorter close to activity maximum (which occurred in the gap between the data sets) than at lower activity periods. However, the amplitudes of the polynomial curves are rather uniform and the intervals of

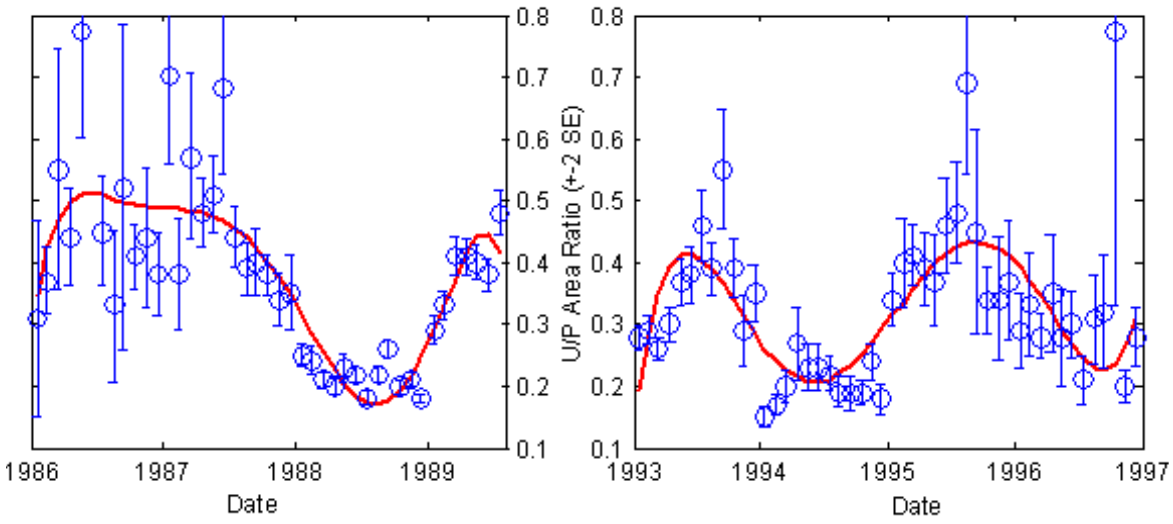

Figure 10: The monthly mean U/P area ratio fitted with polynomials of degree six. The original data points have error-bars corresponding to two standard errors.

available data are short compared to the cycle length of the fluctuation. It is therefore difficult to conclude anything else than that the variation has a period that may vary from 1.5 to 3 years or more depending on how much confidence one has in the data of 1986 and the first half of 1987.

4 Possible Physical Interpretations of the U/P Fluctuation

In this section we will give two phenomenologically different qualitative suggestions of possible physical interpretations of the observed U/P fluctuation, and a semi-empirical quantitative estimation of one of them.

4.1 What Can Affect the U/P Area Ratio?

There are several different parameters that can be involved in a process making the umbra-penumbra ratio of sunspots fluctuate in the way described in Section 3. We will here list some possible ideas and give some suggestions of how they can be investigated in the future.

1) Small sunspots are predicted to be located in a shallow upper layer of the photosphere by monolithic models. Therefore they can be more sensitive to variations in the photospheric pressure stratification, caused by temperature or magnetic field strength fluctuations, than larger sunspots. We will in Section 4.2 present how a fluctuation of the photospheric pressure stratification can give rise to a variation of sunspot hydrostatic equilibrium conditions, and how much a certain pressure level has to change to account for the observed U/P fluctuation.

2) Compactness or complexity of sunspot groups can be varying due to varying conditions in the processes of sunspot group emergence and/or decay. The size of the penumbra associated with a small umbra can depend on the complexity of the magnetic configuration the sunspot is a part of, or in other words how the sunspot is magnetically connected with other magnetic regions. This may actually be possible to get a hint of by further investigations in the DPD. The time behaviour of some kind of compactness parameter, which should by cautiously defined to suit the question at issue, could be examined for all available sunspot groups, or an appropriate subset chosen by some criteria.

4.2 A Zeroth Order Approach

We have performed a simple estimation of how much the effective pressure level has to change in a simple hydrostatic sunspot model to explain the observed phenomenon. The cool parts of a sunspot are very wide and thin, almost of the shape of a coin. The visible umbra and penumbra form a saucer-like depression of the photospheric surface, called the Wilson depression. As an unelaborate approximation we assume that the lower parts of a sunspot can be described by a vertical, cylindrical and circularly symmetrical, magnetic flux tube.

In the upper part of the photosphere, the surrounding gas pressure is decreasing sufficiently fast with height, for the flux tube to fan out to let the internally dominant magnetic pressure decrease at the same rate as the external gas pressure. Hereby hydrostatic equilibrium is maintained. For simplicity we replace the curvature of field lines seen in Figures 2 and 3 in Section 2.3 with a sharp bend at a depth equal to the depth of the Wilson depression of the sunspot. At this depth the field lines of the vertical, cylindrical boundary surface umbra-photosphere beneath the bend is unified with the field lines representing the boundary surface penumbra-photosphere above the bend. This is illustrated roughly in Figure 11. The umbra is in this model represented by the area inside the vertical part of the flux tube, whereas the penumbra area is equal to the projection of the fan-out region on the photospheric surface. Since the fanning out of a flux tube indeed takes place in a

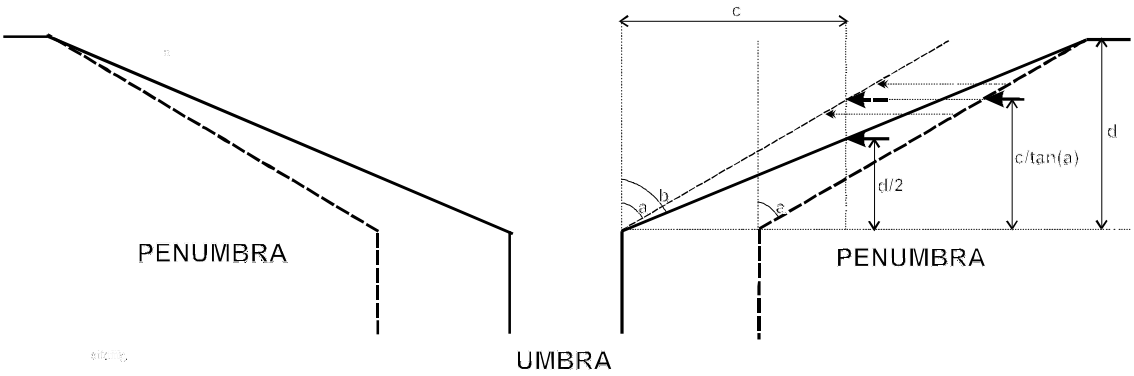

Figure 11: Sketch of the cross-sections of two sunspots with the same total projected area on the photospheric surface, but different U/P area ratios. A sunspot with low U/P is shown with thick solid lines and a sunspot with high U/P with thick dashed lines. The vertical lines in the bottom of the figure represent the cross-section of the umbra-photosphere boundary, the tilted lines represent the penumbra-photosphere boundary and the top horizontal lines the photospheric surface. The right part of the figure is enlarged in Figure 12.

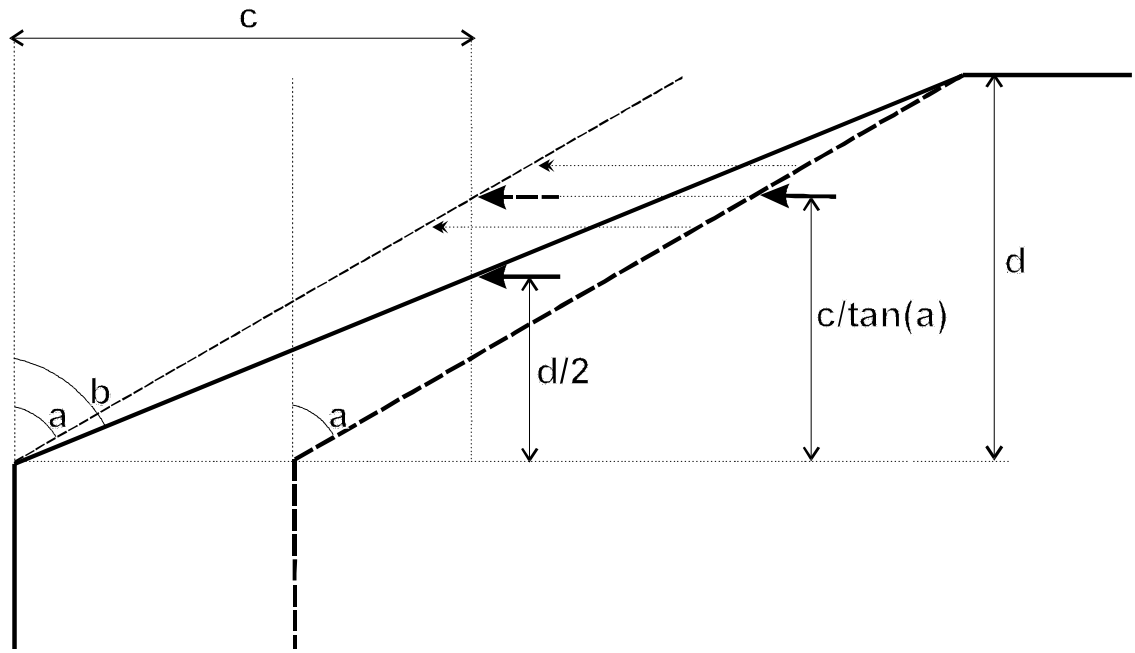

Figure 12: Enlargement of the right part of Figure 11. The angles of the penumbra-photosphere boundary surface with respect to the vertical direction are a for the high U/P ratio spot and b for the low U/P ratio spot. Half the penumbra radial distribution is called c and the depth of the Wilson depression is called d.

shallow layer of the upper photosphere, this linear approximation of the curvature of field lines should be good enough for the rough estimation we are about to make.

A schematic sketch of the cross-sections of two sunspots with different U/P area ratios are shown in Figure 11. The sketch illustrates the model assumptions described above. In Figure 12 the right part of Figure 11, where some parameters used for calculations are illustrated, is enlarged.

The parameters shown in Figure 12 are the angles a and b of the penumbra-photosphere boundary for the high and the low U/P ratio spot respectively, one half of the radial penumbra extension of the low U/P ratio spot, called c, and the depth d of the Wilson depression. The parameter d has the same value for both spots.

According to Collados et al. [Co87], who have made a statistical study of the geometrical Wilson effect, the depth of the Wilson depression for a spot with a diameter of totally 30 000 km is about 400 km. There are no statistical, or other, studies made on whether this depth is of the same order for spots of different sizes or if it scales with spot size.

We assume as an intuitive zeroth order approach, as the monolithic models suggest, that the vertical scaling of a sunspot is proportional to the horizontal scaling. Consequently, the depth of the Wilson depression is proportional to the total size of the spot. Evidence that neither this is the case, nor that it is not, has this far been found. We would like to describe the behaviour of the U/P area ratio with one parameter alone as a simplest approach. To be able to do this, we define the effective pressure level as the level at which the resulting force of the integrated external pressure acts on the penumbra-photosphere boundary.

We have plotted the pressure stratification according to the Harvard-Smithsonian reference atmosphere for the lower chromosphere and upper photosphere, described in [Gi71], in Figure 13. It seems reasonable to approximate a linear pressure stratification in the upper photosphere. The resulting effective pressure level for a linear stratification is located at half the total depth. This will be used in all further calculations.

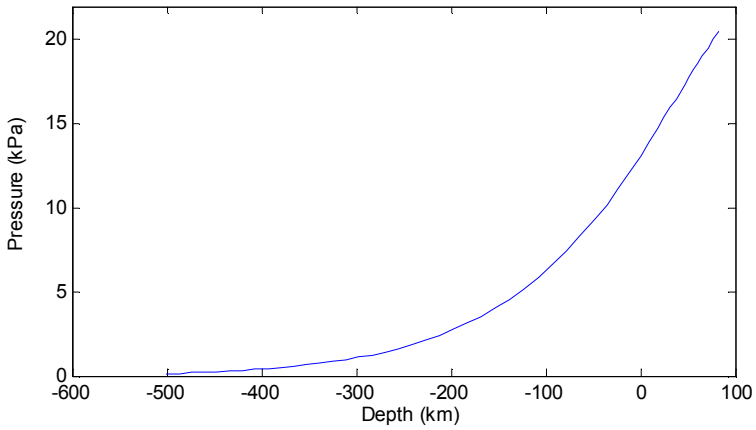

Figure 13: The pressure stratification in the lower chromosphere and upper photosphere of the Harvard-Smithsonian reference atmosphere described in [Gi71]. The boundary between the chromosphere and the photosphere is defined as where the optical depth at a wavelength of 5000 Å is equal to unity. This is the zero level of the solar atmosphere.

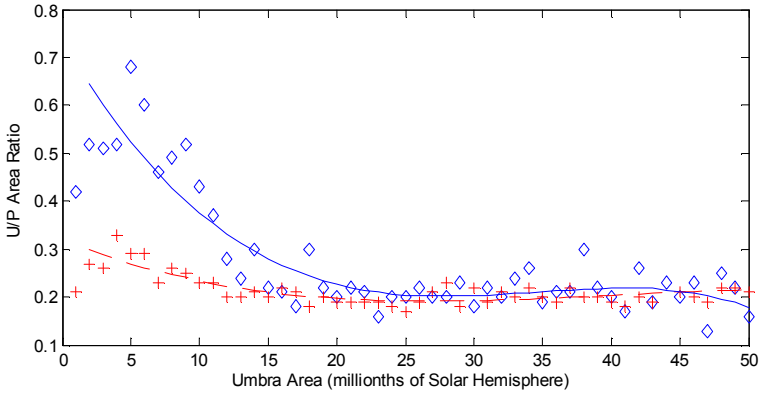

Figure 14: The mean U/P area ratio as a function of umbra area expressed in millionths of the solar hemisphere. The diamonds represent U/P averages of integer umbra areas for the fluctuation maxima of 1993 and 1995. The plus-signs represent U/P averages of integer umbra areas for the fluctuation minima of 1988 and 1994. The solid and dashed lines are least-square fitted polynomial curves of degree three, for which the averages when the umbra area is equal to one unit are neglected.

The mean U/P area ratio as a function of umbra area expressed in millionths of the solar hemisphere is plotted in Figure 14. Here we have treated the years 1993 and 1995, when maxima occurred, and the years 1988 and 1994, when minima occurred, separately. The solid and dashed lines are least-square fitted polynomial curves of degree three, for which the averages when the umbra area is equal to one unit are neglected. It can be seen in the left part of Figure 14 that the spread of the mean values from the fitted curves increases with decreasing umbra area. Both fluctuation maxima and minima have low U/P values corresponding to an umbra area of one unit, compared to the main trend. This may be a sign of a real decrease of the U/P ratio for the very smallest spots, but as a first approach we consider the values too uncertain to trust, since the error introduced due to rounding of areas to integer values is largest for the smallest spots. It should be noted that our assumption here is a bit doubtable, since the U/P values not only for the smallest spots, but for an umbra size less than five units during maxima and less than four units during minima, are below the fitted curves. Different curves than those shown in Figure 14, with global maxima at five and four units can be imagined as fitting the values better.

To estimate how much the effective pressure level has to change to reproduce the U/P fluctuation, we proceed as follows: Consider a sunspot of an arbitrary total size during U/P minimum to have its effective pressure level at half the depth of its Wilson depression. In Figure 12 the height of this level is measured from the bottom of the Wilson depression and labeled $d/2$. During U/P maximum a spot of the same total size has a bigger umbra and a smaller penumbra, but an equal depth d of the Wilson depression. We assume that the internal pressure of the sunspot is equal during maxima and minima, and that a higher U/P ratio, here interpreted as a smaller fan-out angle of the flux tube, is caused by a higher external gas pressure. This is illustrated in Figure 12 by a resulting effective pressure force of the same absolute value as was found at $d/2$ during minimum to be situated at a higher level during maximum. The difference between the levels is found to be

$$c/tan(a) - d/2, \tag{9}$$

by the geometrical and trigonometrical drill illustrated in Figure 12.

Using the fitted curves of Figure 14, describing the U/P ratio as a function of umbra size during maxima and minima, we calculate the umbra radius u, the total spot radius up and the penumbra radial extension p as functions of umbra area U according to

$$u_{\max,\min} = \sqrt{\frac{U}{\pi}}, \tag{10}$$

$$up_{\max,\min} = \sqrt{\frac{U}{\pi} \cdot (1 + \frac{1}{U/P_{\max,\min}})} \text{ and} \tag{11}$$

$$p_{\max,\min} = up_{\max,\min} - u_{\max,\min}. \tag{12}$$

From these quantities, the parameters illustrated in Figure 12 can be calculated as

$$a = \arctan(\frac{p_{\max}}{d}), \tag{13}$$

$$b = \arctan(\frac{p_{\min}}{d}), \tag{14}$$

$$c = \frac{p_{\min}}{2} \text{ and} \tag{15}$$

$$d = \frac{400}{30000} \cdot 2up, \tag{16}$$

where the values of $p_{\max}$ and $p_{\min}$ are chosen from pairs of spots with equal total spot radius up, and by using the assumption of scalable depth of the Wilson depression with 400 km as the depth of a spot with diameter 30 000 km.

The effective pressure level difference, as defined in Equation 9, needed to reproduce the U/P fluctuation, is presented as a function of total sunspot radius by the solid line in Figure 15. For sunspots of small and medium sizes, the level variation is 10-20 km. If we on the contrary to Figure 14 assume that the U/P ratio is constant for all sizes of sunspots during minima, having a value of 0.197, the dashed curve showed in Figure 15 is achieved. The comparison of the two curves illustrates how sensitive the model is for a change of the input values.

4.2.1 Discussion

With the simple estimation technique used in Section 4.2, a relatively small effective height variation is found to be sufficient to reproduce the observed U/P fluctuation. Thus, a variation of effective height should not be excluded as a possible candidate for the physical explanation of the phenomenon. However, a more sophisticated model should be computed in further investigations.

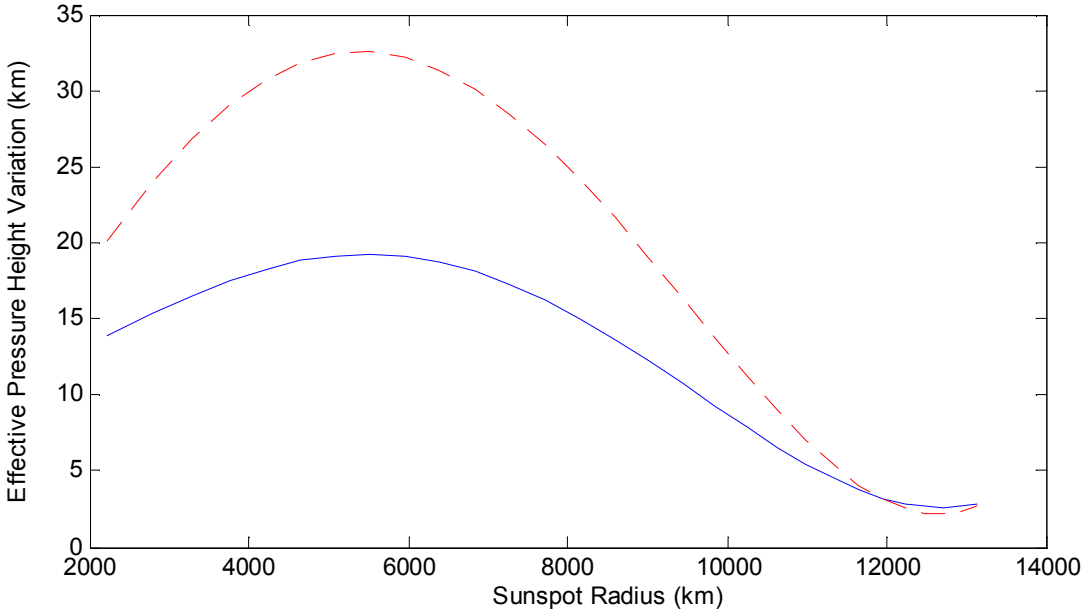

Figure 15: The effective height difference needed to reproduce the U/P fluctuation, as a function of total sunspot radius. The solid line shows the result with the fitted curves of Figure 14 used as input values, whereas the dashed line is produced with a constant value of 0.197 used as U/P input for the minimum period.

5 Mid-Term Periodicities of the Sun

There exist several reports on different solar quasi-biennial oscillations. We have made a survey, mainly to find the main candidates of variations that can have a plausible connection to the U/P area ratio fluctuation.

When relevant data of fluctuations with resembling time-behaviour have been accessible, the same procedures as were carried out on the U/P fluctuation in Section 3.2 have been conducted.

5.1 The Solar Radius

Changes of the solar radius R is in subtle connection with changes in the temperature T and luminosity L of the star. If the luminosity varies with $\Delta L/L \approx 10^{-3}$ the radius and temperature at the solar surface also vary with $\Delta R/R \approx 10^{-5}$ and $\Delta T/T \approx 10^{-4}$ [Cal02]. However, it is not possible to uniquely define a radius for the Sun in view of the fact that stars are gaseous self-gravitating spheres. For the solid planets the radii are defined as a surfaces of constant density, but for the Sun it is usually described as a function of both temperature and density profiles, derived from the drop-off of solar brightness at the limb.

Several ground-based measurement series of the solar diameter have been conducted during the last decades. The earliest written papers of observations of a variability of this parameter are dated already in the first half of the nineteenth century. Due to difficulties of defining the solar limb, systematic effects of atmospheric and instrumental origin, and variations caused by personal bias in the case of visual observations, there is a large disparity of different data series.

5.1.1 Measurements at the Calern Observatory

Laclare et al. [La96] have made measurements of the solar diameter with the solar astrolabe at the Calern Observatory (Observatoire de la Côte d'Azur) in France since 1975. An overview of their measurements as well as data analysis will follow. A more detailed review than given here of how the observations are conducted and the principle of the astrolabe instrument can be found in the references of [La96].

The Calern Observatory astrolabe enables 22 measurements of the solar diameter during a day with optimal meteorological conditions. The zenith distances for the measurements are 30°, 34°, 37.5°, 41.5°, 45°, 49°, 52.5°, 56.5°, 60°, 65° and 70° in both the east and west directions. The observational data series contain a higher number of easterly measurements due to average meteorological conditions with clear mornings and cloudy afternoons. In [La96] several systematic effect are investigated. We have used their data set from 1975 to 1998, which contains date of observation, measured solar semi-diameter, zenith distance and indication of east/west direction.

5.1.2 Zenith Distance and Heliographic Latitude

The thickness of the atmosphere changes noticeably with zenith distance. The mean semi-diameter of the sun decreases with increasing zenith distance in the Calern Observatory data set according to [La96]. The standard deviation, on the other hand, increases with zenith distance. The deviations of the means for specific zenith distances from total average are shown in Table 3. In our analyses of the data set, we have taken the zenith distance effect into account by subtracting the appropriate deviation presented in Table 3

Zenith Angle (degrees)	Deviation (arcsec)
30.0	+0.07
34.0	+0.07
37.5	+0.07
41.5	-0.01
45.0	+0.03
49.0	+0.04
52.5	+0.01
56.5	0.00
60.0	0.00
65.0	-0.03
70.0	-0.09

Table 3: Deviations of the mean semi-diameter from the total mean for each zenith angle. The deviation systematically decreases with increasing distance from zenith, which is called the zenith distance effect [La96].

from each observation. This has also been done in investigations made by the observing team themselves presented in [La96]. In this way a corrected data set with the same average for all zenith distances is produced.

Another interesting feature described in [La96] is a possible diameter dependence of heliographic latitude of the measurement, i.e., with the inclination of the measured diameter to the solar equator. These deviations are somewhat smaller than the deviations presented as a function of zenith distance, and can according to the authors be due to the limited quality of the measurements and their non-simultaneity. If not, this could be a hint of solar oblateness.

5.1.3 The Semi-Diameter Data Series

A plot of all measurements of the semi-diameter shows a large dispersion even on a day-to-day basis. Another feature of the series are frequent gaps unevenly spaced in time and extension due to, e.g., meteorological conditions and technical maintenance of the instrument. As can be seen in Figure 16, there are usually two to three months of total absence of observations during November to February, the winter period of the northern hemisphere where the Calern Observatory is located.

Filtering of the observational data without first averaging over some period gives a very cumbersome result. It therefore seems advisable to first make monthly averages to get a more continuous data set, which is suitable also in the sense that this is the same treatment as the U/P area ratio measurements were given. The years of 1975-1977 have been left out of our investigations due to lack of observations (see Figure 16). The available monthly averages for the period 1978 to 1998 are plotted in Figure 17. These series give a hint of some variability of the observed radius. Not shown in Figure 17 are monthly averages composed by less than three observations: we noticed some very divergent monthly averages with values higher than can be found among the monthly averages published in [La96]. It turned out that all of these were composed of less than three observational values. Laclare et al. have not mentioned disregarded observations in their article, but we have decided to not include these averages in further studies, as

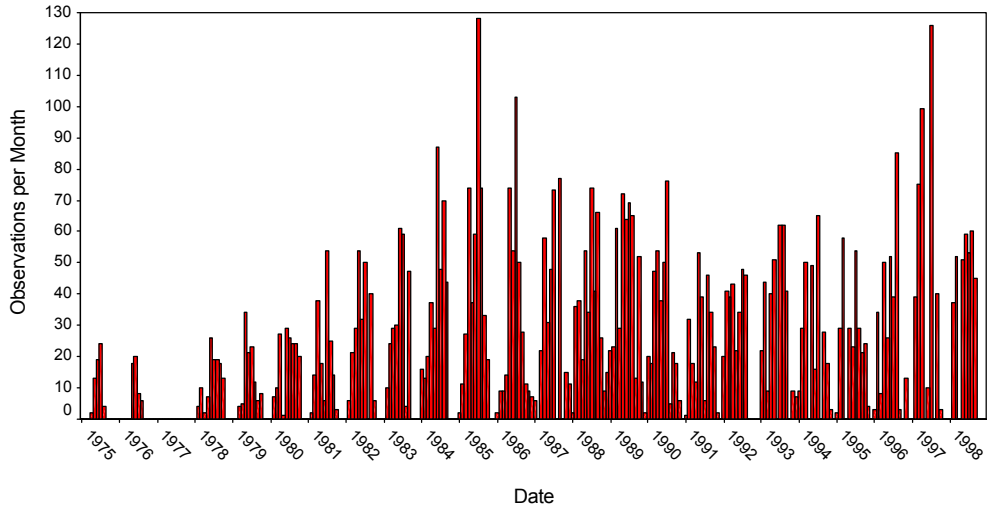

Figure 16: Number of solar diameter measurements per month at the Calern Observatory in France.

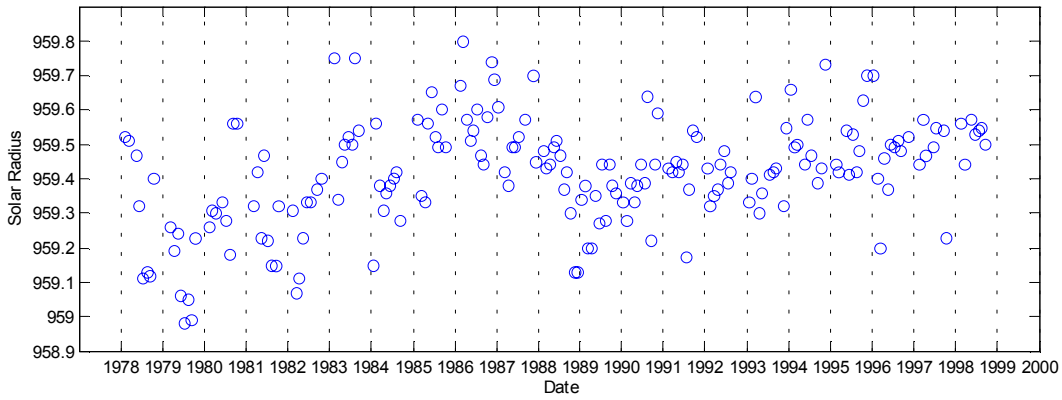

Figure 17: Monthly averages of the solar radius measurements from the Calern Observatory.

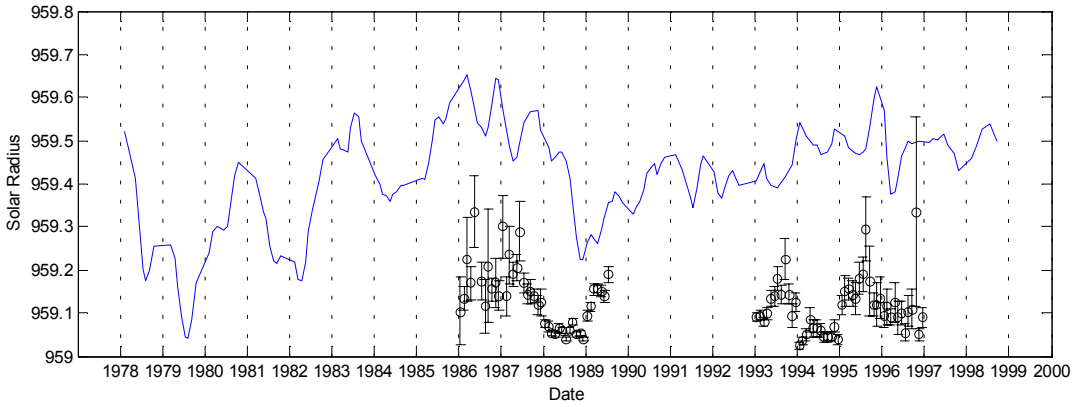

Figure 18: Solar radius variations and the U/P area ratio (in arbitrary scale). The solid line shows a moving average of three consecutive monthly mean values of the solar radius measurements. The U/P area ratio is plotted as monthly means with error-bars corresponding to two standard errors.

they are uncertain but in a filtering process strongly would affect the outcome. When all monthly averages made up of less than three observations are disregarded, a total of 10 monthly averages, evenly spread in the data series, are excluded among the 198 monthly averages available.

5.1.4 Our Analyses

To be able to detect possible periodic behaviour we have filtered the data with a phase-preserving moving average method in Matlab. The most revealing result is achieved by a three-month moving average, which is plotted as a solid line in Figure 18. It should be noted that three-month moving average here means average of three consecutive months for which observations are available. During winter, an average therefore can contain, e.g., November, February and March if December and January are missing. As a comparison, the U/P area ratio fluctuation is plotted in arbitrary scale in the lower part of the same figure.

It is difficult to draw any conclusions from the simple comparison made in Figure 18. The fluctuations of the solar radius measurements may be composed of several parts with different amplitude and different length of cycle. Also, the amplitudes and cycle lengths of potential periodicities may change with time, which complicates the matter. A spectral analysis would be the appropriate way to further investigate possible periods in the data. Unfortunately, the many gaps in the time series make frequency detection by ordinary methods difficult.

Two separate methods of dealing with non-uniformly sampled signals tried on the Calern solar radius data by Moussaoui et al. are described in [Mo01]. In one of these methods, the data are fitted with harmonic functions, whereafter a power spectrum can be obtained with a usual periodogram. In the other method, the data are deconvolved from the observational temporal window by a daily mean uniform sampling and standard FFT routines. The most prominent semi-diameter variation found by both methods are a 10-year periodicity that, as can be seen in Figure 18, has an opposing phase to the solar activity cycle. Minima are found around 1979 and 1989 when solar activity maxima occurred. Two other strong fluctuations have periods of 5.5 and 2.5 years. The three periodicities have amplitudes of the order 0.10", 0.07" and 0.06" respectively, which correspond to 0.10, 0.07 and 0.06 per mille or about 70, 50 and 40 km.

5.1.5 Discussion

A 5.5-year period has not only been found in solar radius observations, but also in, e.g., the solar activity, the geomagnetic field and the period of rotation of the Earth (see [Dj96] and Section 5.3). The 2.5-year period has similarities with even a larger number of mid-term periodicities, and perhaps also the U/P area ratio.

The 11-year period with opposing phase to the solar activity cycle is not only, as can be seen in Figure 18, present in the measurements. It is also present in the *number* of measurements shown in Figure 16. This could be a sign of dependence of average meteorological conditions on the solar activity cycle, if the measurements are conducted with a uniform amount of observing hours per month.

It can be worth mentioning that with the use of full-disk images of the Sun as observed in the neutral iron line at 5250 Å, taken at the Mt. Wilson 150-foot tower, Ulrich and Bertello have made apparent radius measurements for the time interval 1982-1994 [Ul95].

They conclude that the apparent radius of the Sun viewed in this wavelength varies in phase with the solar cycle, in contrast to the results presented above.

5.2 Solar Spectroscopic Temperature

As have been suggested in Section 4.1, a possible interpretation of the U/P area ratio fluctuation could be a varying photospheric temperature. It is, however, a very difficult task to reveal information about variations of the photospheric temperature, since measurements unveil variations of its effective temperature. The effective temperature is, at least in the upper photospheric layers, sensitive to the dynamic distribution of faculae and plage. The temperature distribution of the silent surrounding photosphere itself is therefore complicated to dissolve.

5.2.1 The Neutral Carbon Line

The considered most prominent attempts to measure the temperature of the photosphere without contributions from faculae and plage have been conducted by Gray and Livingston [Gr97a]. They have monitored the neutral carbon line at 5380 Å from 1978 to 1992. This line is emitted from a layer situated deep in the photosphere, coincident with the continuum-forming layers. Gray and Livingston points out the diminishing and almost vanishing of the line at the solar limb as being an observational evidence of its deep formation. After allowance of seasonal and instrumental effects, they have studied the ratios of the carbon spectral line depth to the spectral line depths of iron at 5379 Å and titanium at 5381 Å. The carbon line grows with increasing temperature, while on the contrary the iron line declines and the titanium line is practically constant. Thus, the ratios can be used to estimate variations in temperature for the layer where the lines are emitted. Calibrations of the temperature sensitivity have been made empirically with a sample of sun-like stars.

We have not been able to use the spectroscopic data series ourselves and can only review the result of Gray and Livingston presented in [Gr97a]. When analyzing the ratios, they have found a signal that is in phase with what is found in the irradiance data (presented in Section 5.5), corresponding to the solar activity cycle with a period around 11 years. This signal has a full amplitude (maximum-minimum) equivalent to a temperature variation of 1.5 K.

Also a weak secular term is found by least-squares fitting of a straight line to the data set. The secular term corresponds to a slow rise of the solar temperature by 0.012 $\mathrm{K{\cdot}yr^{-1}}$ deduced from the iron ratio and 0.016 $\mathrm{K{\cdot}yr^{-1}}$ deduced from the titanium ratio. It should nevertheless be pointed out that any secular prediction is very uncertain, taking the relative shortness of the data set (14 years) into account.

5.2.2 Discussion

Most interesting for us are the line depths of the separate lines themselves shown in Figure 19. All three lines contain similar variations in time of a bigger amplitude than the ratios or residuals of any two lines relative to each other. These variations seem to have a quasi-biennial behaviour that varies in phase, but with different amplitude. The amplitudes are smaller for weaker lines but do not scale with the depth of the lines indicated on the vertical axis of Figure 19. Gray and Livingston interpret these variations not to arise from temperature fluctuations since the excitation and ionization processes

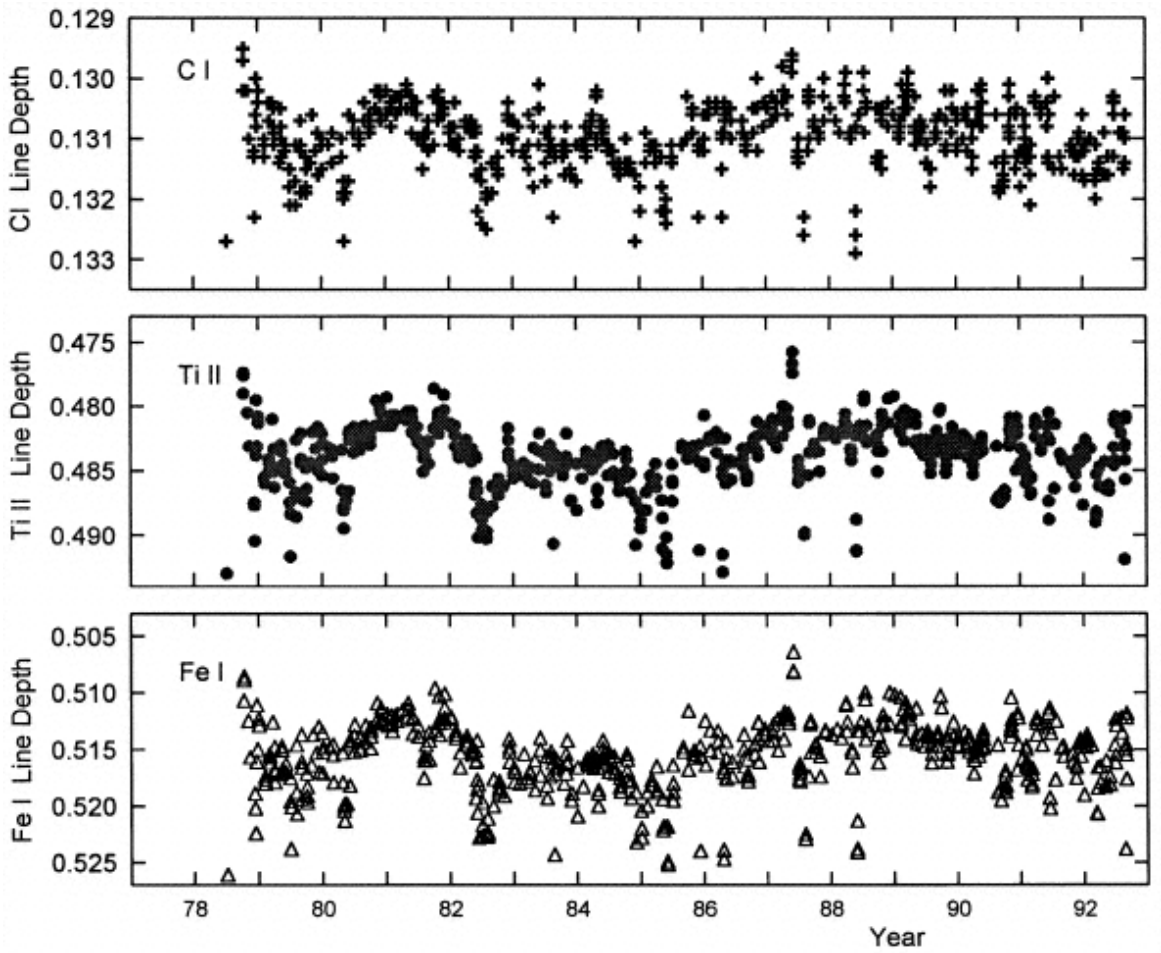

Figure 19: Line depths for C I λ5380 in the top panel and its adjacent lines Ti II λ5381 and Fe I λ5379 in the two lower panels. An essentially identical pattern is seen in each panel [Gr97a].

of the three species are sufficiently different to cause real temperature fluctuations to show up as different, and not similar, variations in the lines. The authors are trying to simulate the observations by a scattered-light hypothesis, but cannot explain how the necessary variation of the scattered light can practically arise in the 13.5 m double-pass spectrometer at the McMath-Pierce telescope on Kitt Peak used for the observations. In a later published errata, it is found that the scattered-light calculations are inconsistent [Gr97b]. This implies that no possible explanation is obtained for the nearly synchronized variation of the central depths of all three spectral lines.

In our opinion the real nature of photospheric temperature variations is not yet clear even for the layer where the spectral lines discussed in this section are emitted and a quasi-biennial fluctuation can not be excluded. Several other spectral lines, and ratios between them, that are possible candidates to use as measurements of temperature exist, but are not yet investigated.

Some criticism of using line-depth ratios as a measure of temperature without carefully taking gravity effects into account, as Gray and Livingston have done, has been presented by Caccin and Penza [Ca02]. Their skepticism lies in the use of sun-like stars for calibration of the line-ratio sensitivity to temperature variations. This usage implies first of all, that differences of the atmospheric stratification of stars along the stellar sequence well

simulate what is happening in the atmosphere of the Sun during its activity cycle. Secondly, it assumes that the ratios do not depend on gravitational forces, since the sun-like stars used for calibration show a significant disparity in their estimated surface gravity, while in the solar case the value of g is practically constant throughout the activity cycle. When Caccin and Penza are taking the latter into account according to a model described in their paper ([Ca02] and references therein), they obtain a photospheric temperature variation of 3 K. This is not only twice as large as was found by Gray and Livingston, but also higher than predicted by irradiance variations.

Caccin and Penza present an alternative model in which there is a variation of the photospheric granulation size of 5%, granular dimensions being minimum at sunspot maximum. They argue that this gives a variation of the line-ratios that is opposite, but of the same order, as a temperature fluctuation gives. In this case a photospheric background temperature fluctuation of 0.8 K is enough for the observed line-depth ratio variations and the similarity in pattern of the individual lines in Figure 19 is a direct consequence of granular size variations. The apparent quasi-biennial behaviour of the patterns, and if this should be interpreted as if the granular size distribution is fluctuating in a quasi-biennial manner, is not commented by the authors.

5.3 Solar Activity

There are several different ways to define and measure solar activity. The most classical indicator is based on counting the number of sunspots on the visible solar hemisphere. Today, much more sophisticated measurements of solar activity are made routinely, but none has the link with the past that sunspot numbers have.

5.3.1 Quasi-Biennial Fluctuations in Solar Activity Indices

There are several reports on quasi-biennial oscillations of different indices of solar activity. The time behaviour of the number of Hα-flares, the index B2 of photospheric magnetic field energy[1] and the total sunspot area during solar cycles 21 and 22 are investigated by Bazilevskaya et al. [Ba00]. Superposed on the general 11-year period they have found correlated quasi-biennial signals in all three indices. The signals are more prominent after separation into northern and southern hemisphere. The opposite hemispheres have oscillations with similar time behaviour, but sometimes shifted in phase with respect to each other.

Quasi-biennial oscillations have also been reported for, e.g., sunspot numbers (Section 5.3.4), green corona activity (Section 5.3.6), solar radio flux at 10.7 cm and Ca^+ K plages ([Dj93] and references therein).

Suggestions have recently been put forward that quasi-biennial oscillations are the main property of the fine structure of the solar activity cycle, which is interpreted as a superposition of a triad of consecutive wave trains appearing with 10- to 11-year intervals [Ko02].

[1] B2 = $\left\langle \frac{B^2}{8\pi} \right\rangle$

5.3.2 Relative Sunspot Numbers

There exist an almost uncountable number of reports on analyses of relative sunspot numbers R, defined as

$$R = K \cdot (10m + n), \tag{17}$$

where K is an observer correction factor (typically 0.6), m the observed number of groups and n the observed number of sunspots. Equation 17 was introduced in 1848 by the Swiss astronomer Johann Rudolph Wolf as a measure of solar activity. R is therefore also referred to as the Wolf Number. The 11-year cyclic behaviour of solar activity was discovered previous to the introduction of the Wolf Number, in 1843 by Heinrich Schwabe.

5.3.3 Extrapolations into the Past

The Wolf Number, as an index of solar activity, has been evaluated for earlier times than 1848 by applying it to older sunspot records. E.g., records of auroral activity, tree rings and ice cores have been used to extrapolate the level of solar activity longer than reliable sunspot records are available ([Pr84] and [La00]). These extrapolations, in combination with the almost total lack of observed sunspots during the second half of the 17th century, even though systematical investigations were made already half a century earlier by Galileo Galilei with the newly invented telescope, make out observational support for big scientific discoveries like the Maunder minimum of solar activity from 1645 to 1715.

According to reconstructions, the signs of the solar activity have not only been disappearing during some intervals, the cycle length may also have been varying much more than during the last well-known 140 years. Between 1750 and 1958 the range has been from 7.3 to 17.1 years with an average of 10.9 years.

5.3.4 Our Analyses

We have in our analyses used the monthly means from 1848 to 2002 of the International Sunspot Number defined by Equation 17. Each daily international number is calculated as a weighted average of measurements from a contributing network of cooperating observatories. International numbers from 1749 onward are available through NOAA NGDC[2], but prior to 1848 the observational series used for the monthly means are not continuous [Ho98]. Before 1818 the number of missing days is great and no observations are available for many whole months. From 1818 to 1848 daily gaps still occur, but less frequently, a few months are missing in total.

The monthly means of the relative sunspot number is plotted in Figure 20 as a thin line for the 30-year interval 1970 to 1999. To be able to detect quasi-biennial variations, a 5-month moving average has been obtained with Matlab, plotted as a thick line together with the monthly averages in Figure 20. Shown in the figure is also a dashed line constructed by a 20-month moving average. This serves as an indicator of the mean trend of the cycle, only varying according to the well-known 11-year period. During some intervals of the solar cycle the 5-month moving average shows a quasi-biennial behavior superimposed on the mean trend, which we intend to investigate further.

In Figure 21 the monthly deviations of the relative sunspot number 5-month moving average from the mean trend are shown as a solid line. In this figure a quasi-biennial

[2] the National Geophysical Data Center of the National Oceanic & Atmospheric Administration, USA

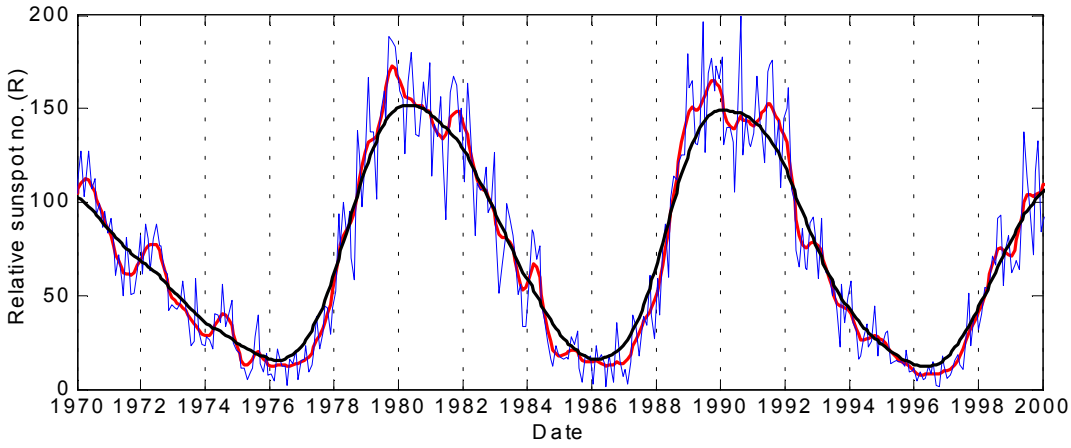

Figure 20: The monthly mean values of the relative sunspot number (blue line), a 5-month moving average (red line) and the mean trend illustrated by a 20-month moving average (black line).

signal clearly can be seen for some periods, but not continuously. The behaviour is similar for other intervals than the 30-year interval shown in Figure 21, suggesting that a quasi-biennial oscillation builds up in the beginning of a solar cycle, reaches its maximal development during solar cycle maximum and decays at the end of the cycle. This is also concluded by Apostolov [Ap85a] in a previous investigation. The correlation of development and decay with different phases of the activity cycle during 1970-1999 can be identified by comparing the residuals of Figure 21 with the mean trend of the relative sunspot number in Figure 20, serving as an indicator of the state of the activity cycle.

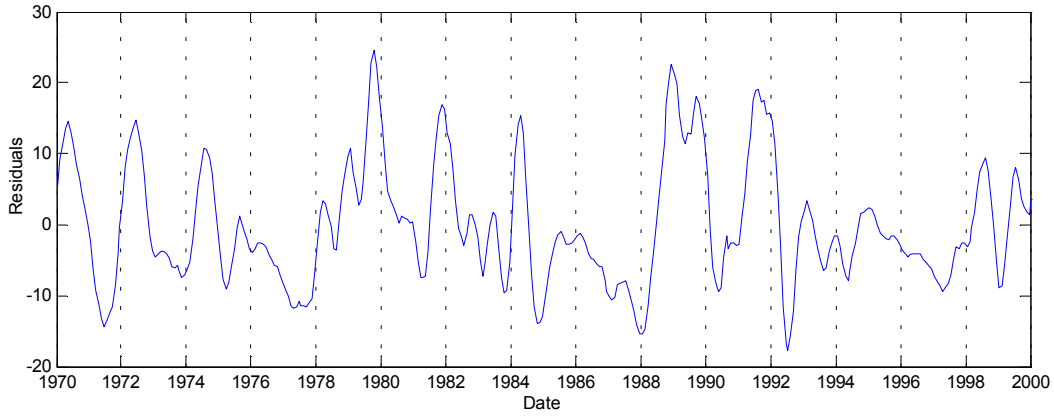

Figure 21: The deviation of the relative sunspot numbers from the mean trend, illustrated by the residuals of 5- and 20-month moving averages.

A spectral analysis of residuals of the mean trend subtracted from the original data for the whole interval from 1848 to the middle of 2002 is shown in Figure 22. The pre-treatment both eliminates and distracts some information about the original data, but normalizes the spectral peak of the high-amplitude 11-year periodicity of the original series and reveals the behaviour of the residuals themselves and the quasi-biennial period. Still, the spectral peak labeled with an a corresponding to the 10.7-year period is the strongest in Figure 22. It is followed in strength by three peaks, denoted b, c, and d, all of almost the same order of magnitude at corresponding periods of 5.2 years, 25.6

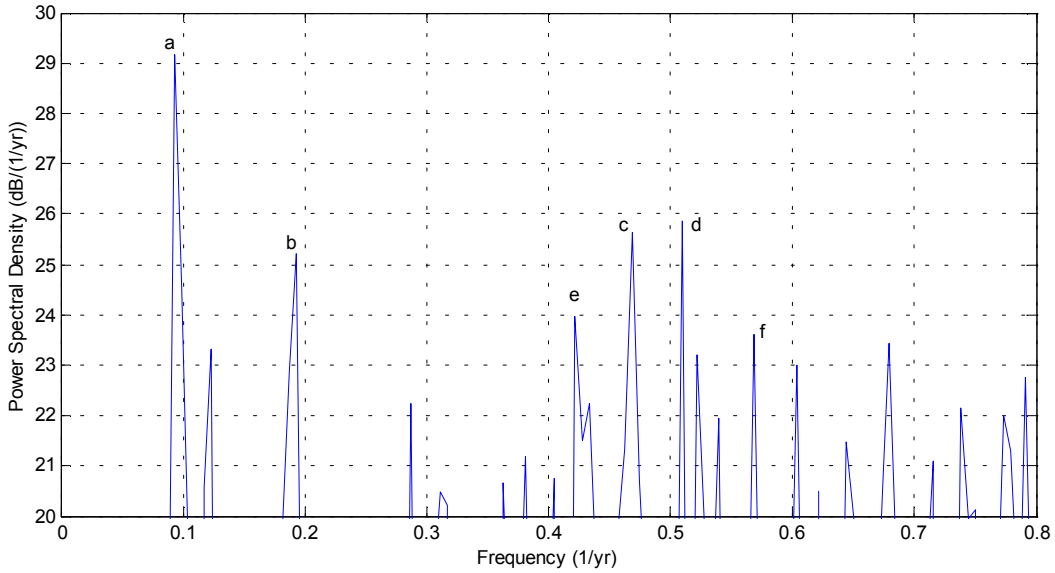

Figure 22: Spectral analysis of the monthly mean relative sunspot number from which we have subtracted the mean trend. The corresponding periods of the labeled peaks are: a = 10.7 years; b = 5.2 years; c = 25.6 months; d = 23.6 months; e = 28.4 months; f = 21.1 months.

months and 23.6 months respectively. There are also a few other, lower, peaks with quasi-biennial periods as well. Most prominent of those are a low-frequency peak labeled e at 0.423 cycles per year having a corresponding period of 28.4 months and a high-frequency peak denoted f at 0.5684 cycles per year corresponding to a 21.1-month period. Apostolov [Ap85a] comments that peaks of these kinds may not necessarily correspond to independent oscillations, but instead are indicators of a variability in the cycle length of only one oscillation. This could result in the, presumably, developing and decaying signal seen in Figure 21.

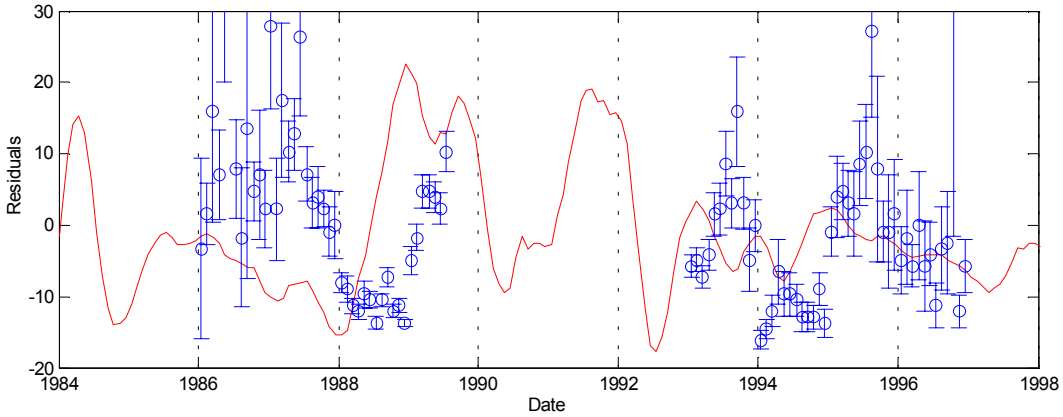

Figure 23: The relationship between the variation of the U/P area ratio found in the DPD and the deviation of the solar activity from its mean trend (solid line). The monthly means of the U/P area ratio are plotted as open circles with error-bars corresponding to two standard errors.

The periods found are close to the ones found in, e.g., solar neutrino data (Section 5.4) and the solar radius observations at the Calern Observatory (Section 5.1). There is also a resemblance with the U/P area ratio periodicity.

The relationship between the variation of the U/P area ratio and the quasi-biennial deviation of solar activity from the mean trend is examined quantitatively in Figure 23. The monthly means of the U/P ratio are plotted as open circles with error-bars corresponding to two standard errors whereas the deviation of the relative sunspot number 5-month moving average from the mean trend is shown as a solid line.

The available time intervals of sunspot data in the DPD are limited to the ascending and descending phases of solar cycle number 22. The quasi-biennial oscillation of the solar activity is on the other hand only clearly detectable during the maximum of the same cycle, but we can suspect a broad but low quasi-biennial maximum with its centre at the very beginning of 1986, an ascending phase during 1985 and slow descending during 1986 and 1987. There may also be local maxima during the descending phase of solar cycle 22 in the very beginning of 1993 and 1995. If we compare the two periodicities shown in Figure 23 it seems likely that an analogy is present, but the U/P fluctuation lags the quasi-biennial variations of the solar activity by possibly 5-7 months. This is supported by the comparison of the U/P data of 1989 closest to activity maximum with the ascending phase of the strong local double-peak of solar activity. A lag of the same order is found also here.

It is not clear how to interpret what causes solar activity, and therefore a plausible connection between activity and the U/P ratio is difficult to establish. Since a 10- to 12-year signal is not visible in the U/P ratio but is the dominant feature of the solar activity, with an amplitude exceeding its quasi-biennial features with several orders of magnitude, the two quasi-biennial signals are most likely different messengers of an underlying mechanism affecting them both if any association is present. This can not be established at the moment, since the photoheliographic data set available is rather limited in perspective of time compared with the lengths of the cycles investigated. If more reliable signs of association can be established is for the future to tell.

5.3.5 The Sunspot Group Number

Sunspot group number values were derived by Hoyt and Schatten [Ho98] to provide an alternative homogeneous record of solar activity from 1610 to 1995. The long-term changes are more self-consistent than the changes deduced using relative sunspot numbers are. The Group Sunspot Number (R_g) is defined by

$$R_g = \frac{12.08}{N} \sum k_i' \cdot G_i, \tag{18}$$

where 12.08 is a normalization constant chosen to make the mean group number identical to the mean relative sunspot number for the interval 1874 to 1976. N is the number of contributing observers, G_i the number of sunspot groups as provided by the i^{th} observer and k_i' the correction factor for the i^{th} observer.

The Group Sunspot Number is considered to contain less observer noise, and the longest cycle has a period of 15 years compared to the anomalously more than 17-year long cycle deduced from the relative sunspot numbers. The differences are biggest for the values before 1848, when the relative sunspot number series are partly, and as claimed by Hoyt and Schatten, badly, reconstructed.

Frick et al. [Fr97] have made wavelet analysis on the Group Sunspot Number series to investigate the long-term behaviour of the main 11-year signal and the Gleissberg cycle, which has a period of ∼100 years. Since they were interested in long time-scales, the quasi-biennial behaviour is not discussed.

Our intentions have been to investigate the Group Sunspot Number series in a similar manner as was done in Section 5.3.4 for the International Relative Sunspot Number, but this has not been done due to lack of time.

5.3.6 The Green Corona Activity

Among others, Apostolov and Letfus have made investigations on the measurements of the green corona intensity at 5303 Å [Ap85b]. The investigations give indications of quasi-biennial oscillations in coronal activity. The variations are found to be similar to the quasi-biennial fluctuations of the solar activity, as measured by relative sunspot numbers, described in Section 5.3.4.

Coronal oscillations depend somewhat on latitude, but for all latitudes the amplitudes are smallest during solar cycle minima and largest around and after cycle maxima as the relative sunspot number fluctuations. The largest amplitudes are observed between latitudes of 10° and 30° on both hemispheres, i.e., where sunspots mostly occur. The characteristic period range is 22-27 months, somewhat varying with latitude, with a mean value of 24 months. The oscillations are lagging the solar activity with approximately 3 months.

Apostolov and Letfus suggest that the green corona oscillations are related to the solar activity oscillations by being caused by a disturbance propagating up from the subphotosphere through the upper solar atmosphere. As the height of the layer of green corona emission is situated about 43 500 km above the photosphere, the velocity of propagation in the case of a three-month lag would be 5 m/s.

5.4 Solar Neutrino Measurements

When the measurement series in the hunt for solar neutrinos were started deep down in the Homestake mine in South Dakota, USA, during the sixties, Raymond Davis Jr., main responsible of the experiment was almost laughed at. Could it really be possible to sort out a few atoms of Argon, produced by the reaction

$$\upsilon + {}^{37}Cl \rightarrow {}^{37}Ar + e^{-}, \tag{19}$$

in a tank filled with 615 ton chlorine in the form of C_2Cl_4, with high enough accuracy? Davis, though, persistently continued the measurements through the years, being able to count on average half a neutrino reaction per day, and the experiment was soon found successful. Actually successful enough for rewarding Davis with the Noble Prize in physics 2002 for his struggles, shared with Masatoshi Koshiba and Riccardo Giacconi.

Many mysteries have been revealed by the Homestake and other solar neutrino experiments. The mysterious missing solar neutrinos, when comparing the capture rates of the measurements to predictions according to the otherwise well working solar model, has achieved greatest attention. This controversy is now considered explained by neutrinos changing flavour, or in other words appearing in different form than expected when reaching the detectors on Earth eight minutes after leaving the solar core. Therefore they are not being captured in a as high as expected rate by experiments only sensitive to the electron-type neutrinos.

Another amusement disclosed is a quasi-biennial variability in the measurements, to be discussed below.

5.4.1 Variations in the Solar Neutrino Measurements

Sakurai reported in the year of 1979 about a quasi-biennial variation of the solar neutrino flux from 1970 to 1976, as measured by Davis et al. in the Homestake experiment [Sa79]. In the article Sakurai compares the observed flux of solar neutrinos with solar activity deviations from the mean trend, as manifested by relative sunspot numbers, for the same interval of time. The method of investigation of the solar activity that Sakurai uses is practically the same as the method we have used on the relative sunspots number series in Section 5.3.4. His results are the same as our, shown in Figure 21, Section 5.3.4. The residuals of the relative sunspot numbers, when the mean trend is subtracted, contain a clear quasi-biennial signal during the interval 1970-1976. Sakurai finds a high correlation between the two quantities, and that the solar neutrino flux is lagging the solar activity residuals with approximately four months. The oscillation is further described by Sakurai in [Sa81] and later on by others as well, e.g., by Liritzis in 1995 [Li95]. We have not been able to perform any analyses of our own on solar neutrino data of any kind and will therefore only shortly review the most recent and ambitious analyses we have found in the literature.

In [Li95] Liritzis discusses all different periodicities found in the solar neutrino data for the interval 1970-1990. The methods of investigation are Fast Fourier Transform (FFT), Maximum Entropy Spectrum Analysis (MESA) and Power Spectrum Analysis (PSA) employing a Blackman-Tukey window. Since the data are unequally spaced, having $\sim$65% blank entries after arranging it into monthly means, they first use a 4-month equally-spaced smoothing sequence and a cubic spline polynomial approximation.

Liritzis classifies the periodicities found into periods longer and shorter than 5 years. There are two periodicities with approximate periods of 10 and 5 years that both can be associated with the solar activity cycle. These periods are present also in our investigations of the relative sunspot number, as peaks a and b in the spectral analysis shown in Figure 22, Section 5.3.4. Liritzis also finds two periodicities with periods shorter than 5 years, one at 1.5 years and one at 2.3 years. The 1.5-year periodicity is strongly variable with time. The highest spectral powers are achieved for the periods shorter than 5 years.

There is a fundamental difficulty in interpreting the results of the search for periodicities in the solar neutrino flux: the restricted quantity and quality of the available data, since the reaction rates of the experiments are very low. Sturrock et al. have made simulations and used different statistical tests on the Homestake data to find evidence for a non-constant neutrino flux from the Sun [St97]. The tests give very different results, but the authors conclude that unless some tests are invalid, it seems likely that the neutrino flux as measured by the Homestake experiment has varied, not only with a quasi-biennial period (2.2 years) but also with the so-called Rieger period of 157 days.

The Sudbury Neutrino Observatory at 2073 meters depth in the Creighton mine, Ontario, Canada has the sensitivity required to determine the existence of neutrino oscillations and can provide evidence for what happens with the solar neutrinos on their way to Earth. With this level of sensitivity the experiment should also be capable of identifying the presence of a quasi-biennial periodicity in the solar neutrino flux if it exists.

5.5 Solar Irradiance

The Sun is radiating $3.853 \cdot 10^{26}$ J of energy into surrounding space every second. This almost incomprehensibly high amount is achieved by the conversion of 700 000 000 tons of hydrogen to 695 000 000 tons of helium during the fusion processes in the solar core.

5.5.1 On the Constancy of the Solar Constant

The solar constant is not a constant at all. This statement is based on the exploration of solar total irradiance records from five independent satellite experiments, somewhat overlapping in time. The overlap enables adjustment and calibration of the different data series to produce a composite data set extending from late 1978 onwards. The needs for adjustment and calibration are due to differences between the data sets, introduced by, e.g., an uncertainty of $\pm 0.2\%$ of the absolute levels of the satellite radiometers and sensitivity drifting due to long-term exposure to the space environment. For these types of detectors the high-energy solar radiation and bombarding particles, which alter the interior surfaces of the receiver cavities and their electronics, are of greatest concern [Fr98].

Before the early 1980s it was not possible to deduce whether the solar irradiance was constant or not, since no variations could be reliably detected from beneath the Earth atmosphere and long-term space borne measurements were not yet conducted.

5.5.2 Monitoring Satellites

The first and hitherto longest lasting of the high-precision space borne experiments monitoring the total solar irradiance was the radiometer HF of the Earth Radiation Budget investigation (ERB) on board the Nimbus 7 satellite between November 1978 and January 1993. It has been possible to compare its data with what was retrieved by the Active Cavity Radiometer Irradiance Monitoring I (ACRIM I) from February 1980 to November 1989, on board the Solar Maximum Mission (SMM).

The Solar Monitor Measurements of the Earth Radiation Budget Satellite (ERBS) started in October 1984 and are taken into account in the composite record until June 1996. ERBS data are well correlated with what were received from the ACRIM II experiment on board the Upper Atmosphere Research Satellite (UARS), from which measurements are used from September 1991 to January 1996. Since the launch and positioning of the Solar and Heliospheric Observatory (SOHO), irradiance data have been provided by the Variability of solar IRradiance and Gravity Oscillations investigation (VIRGO). SOHO was launched in February 1996 and is situated at the inner Lagrangian point (L1), 1.5 billion meters out in space, or about one percent of the way to the Sun as seen from Earth. Irradiance data have been received ever since, except for the interval June 25 to October 7 1998 when SOHO was lost in space. This period is distinguishable as a gap in the composite irradiance series.

5.5.3 About the Variability

Variations of the measurable solar irradiance might originate from at least three other causes than the blocking by sunspot regions and enhancements due to bright faculae ([Gr97a] and references therein): changes of the solar effective temperature, changes of the solar radius or an irradiance dependence on latitude that changes in time. The latter can, very simplified, be interpreted as a non-isotropic output of solar power where the

angular dependence of the emission varies during the activity cycle. The Earth and SOHO with its VIRGO experiment would therefore receive a higher irradiance during some periods than other, due to beaming of solar emission.

Periodicities with cycle lengths of the order of a solar rotation or shorter are usually directly associated with flux redistribution from photospheric magnetic fields. Big, long-term variations are attributed to the changing emission due to bright faculae and the magnetic network. Speculations suggest that a varying global brightness component may be present, such as a few-degree change in solar surface temperature, in addition to the magnetic sources [Le97]. Simulated effects of sunspot darkening and facular brightening only, can not completely explain the total irradiance variability [Fr98].

The global mean surface temperature of the Earth increased by 0.2°C during 1986-1996. This can not be explained by the measurements of total irradiance, which after correction for a variety of instrumental effects show comparable levels at the two minima in the beginning and the end of the interval.

5.5.4 Our Analyses

We have made some investigations on a total solar irradiance composite data set produced by PMOD/WRC[3] from measurements by the satellite experiments mentioned in Section 5.5.2. Details about the data set can be found in, e.g., [Fr98]. The monthly means of the irradiance composite data set and the mean trend disclosed by a 20-month moving average are illustrated in Figure 24. Variations shorter than one solar rotation, which can be regarded as noise as they are caused by the appearance and disappearance of active regions on the visible solar hemisphere, are smoothed out by the use of monthly means. Also, monthly means eliminate the otherwise every now and then appearing gaps in the daily data series to leave us with only one gap, occurring July to September 1998 when SOHO was unable to perform observations.

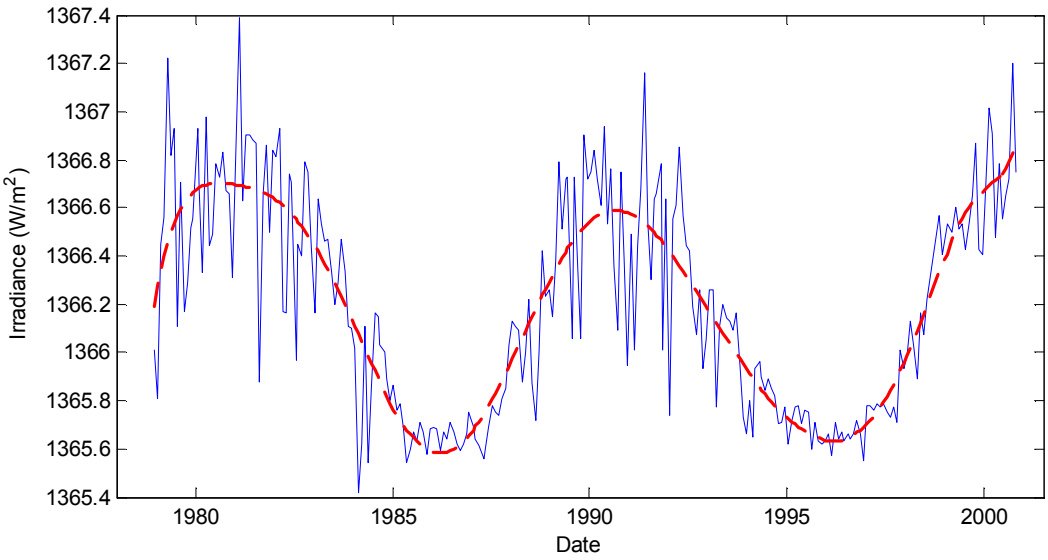

Figure 24: Monthly mean of the total solar irradiance composite data set (thin line) and the mean trend illustrated by a 20-month moving average (thick dashed line).

[3] Weltstrahlungszentrum, Physikalisch-Meteorologisches Observatorium Davos, Switzerland

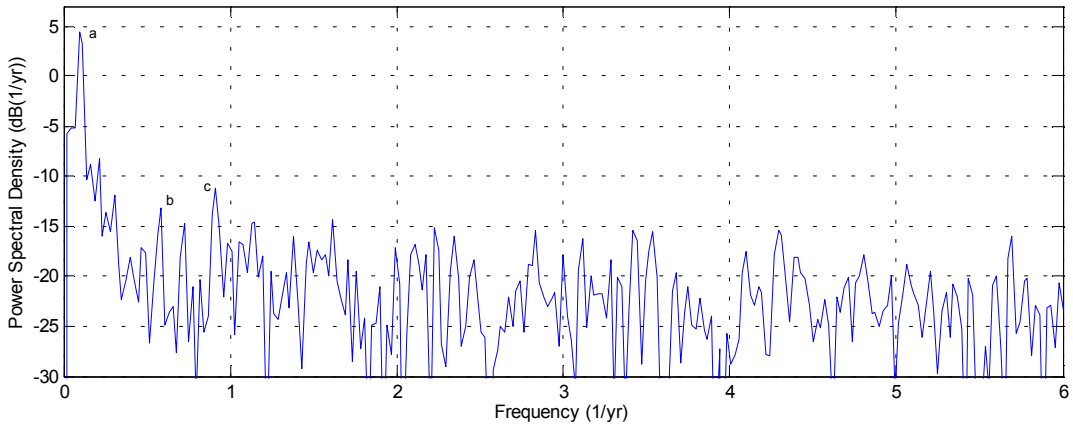

Figure 25: Power spectral analysis of the monthly means of the composite total solar irradiance data. We have transformed the data to a fluctuation around zero by subtracting the mean value 1366.18498 W/m^2 prior to analysis. The labeled spectral peaks correspond to periods of length a = 10.6 years, b = 20.5 months and c = 13.1 months.

Prior to the spectral analysis shown in Figure 25, we have used linear interpolation to fill the gap of the missing months and produce a continuous time series. We have also transformed the data to a fluctuation around zero by subtracting the mean value of 1366.18498 W/m^2. The spectral peaks in Figure 25 have periods corresponding to a = 10.6 years, b = 20.5 months and c = 13.1 months. As the peak b suggests a quasi-biennial period hidden in the data set, we have tried different filtering methods to extract the temporal behaviour. Shown in Figure 26 are the residuals of monthly total irradiance data minus the mean trend and a 5-month moving average of this parameter. Quasi-biennial behaviour with the same phase, but slightly different amplitude and containing a superposition of some higher frequency oscillations is found when shorter moving-average lengths are used. A similar pattern turns up from least-square polynomial approximations of different degrees as well.

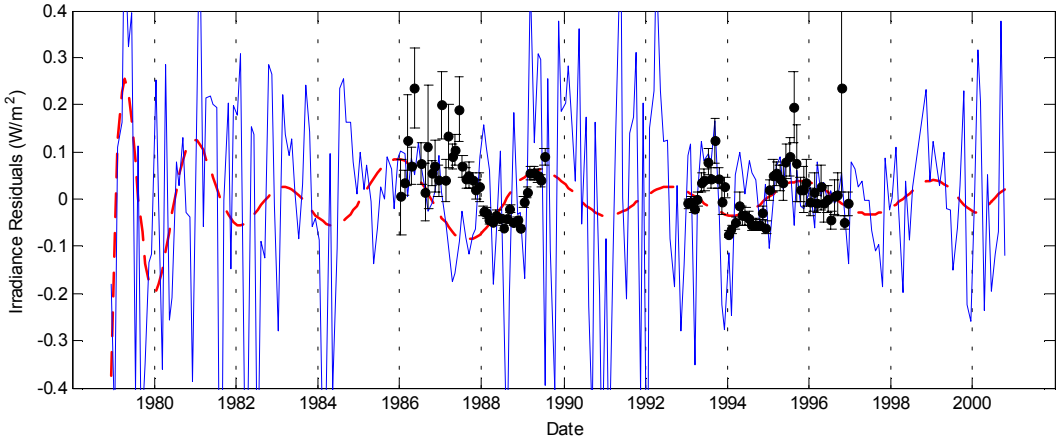

Figure 26: Residuals of monthly mean total solar irradiance minus the mean trend (thin line) and a 5-month moving average of this parameter (thick line). The U/P fluctuation is plotted in arbitrary scale as monthly means with error-bars corresponding to two standard errors.

To be able to compare the possible quasi-biennial signal of the irradiance data with the U/P area ratio fluctuation, the latter is plotted in arbitrary scale as monthly means with error-bars corresponding to two standard errors in Figure 26. The U/P fluctuation seems to lag the irradiance variation, except at their simultaneous maxima of 1996 when the two periodicities are in phase. This inconsistency, if the found quasi-biennial signal in the irradiance data is real, may not necessarily contradict a possible connection between the two quantities compared. The inconsistency can instead be caused by the rough method used to disclose the possible quasi-biennial behaviour of the irradiance data. More sophisticated methods should be used to resolve these matters in further studies.

5.6 The Sun as Seen in Ultra-Violet

As shown in Section 5.5 the total solar irradiance is not constant, but the variations are only 0.10-0.15%, an immeasurable quantity from beneath the Earth's atmosphere. Solar ultra-violet (UV) irradiance between 150-400 nm makes out less than 8% of the total irradiance, but contributes with up to 30% of the overall variations. Sunspots emit a smaller fraction of their radiation in the UV-range than the surrounding photosphere. On the contrary the hot faculae emit relatively more UV radiation than the photosphere does.

5.6.1 The Mg II Index

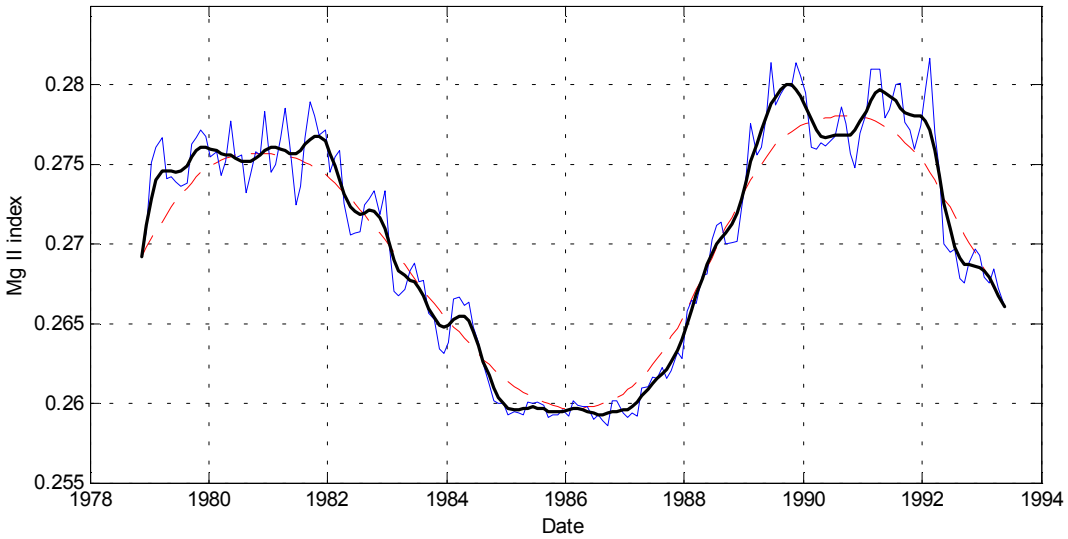

Figure 27: Monthly means of the Mg II index (thin line), the mean trend produced by a 20-month moving average (dashed line) and a 5-month moving average showing quasi-biennial behaviour (thick line).

The Mg II index is a daily measurement of solar UV variability using the Mg II absorption feature at 280 nm, derived by Heath and Schlesinger [He86] and shown in Figure 27. This index can with different scaling factors estimate solar irradiance variations at any wavelength in the 170-400 nm wavelength range, and is used as a dimensionless quantity for measurements of solar chromospheric activity.

The Mg II index is constructed by calculating a ratio with the average of three consecutive wavelength samples at the core of the Mg II H and K doublet at 280 nm in the nominator and the average of two pairs of consecutive wavelengths equally spaced, about 3 nm from the core wavelength in the denominator. The index was first constructed using measurements from the Solar Back-scatter Ultra-Violet instrument (SBUV) on board Nimbus 7, but has been extended to the NOAA-9 and NOAA-11 SBUV/2 instruments under the name composite Mg II index.

5.6.2 Our Analyses

We have investigated the composite Mg II index, available for the interval late 1978 to early 1993 through NOAA NGDC, in a similar manner as the total solar irradiance data in Section 5.5.4. As for all observational data sets we have used, also this sequence contains gaps. To achieve a continuous time line for spectral analysis and for the reasons of equal treatment we have produced monthly means, which are used in all analyses. Among the monthly means, still two values are missing, October 1988 and March 1991. We have filled these gaps with help of linear interpolation of neighbouring values prior to spectral analysis.

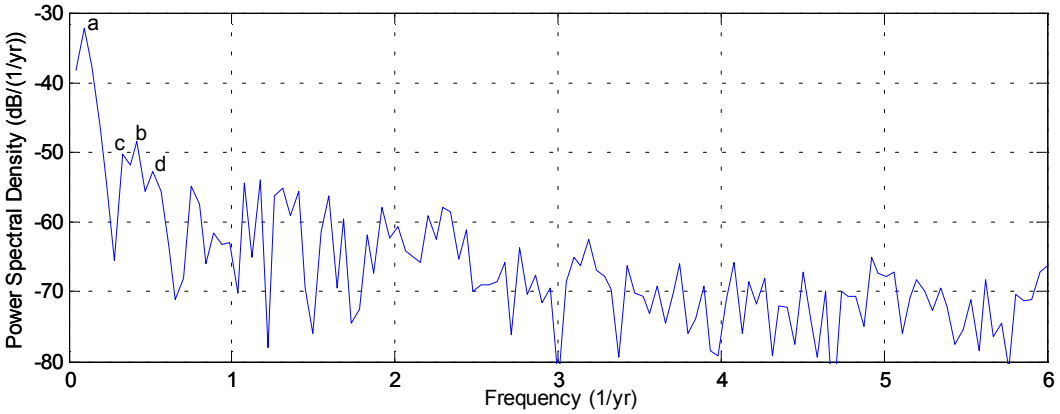

Figure 28: Spectral analysis of the Mg II index, November 1978 to May 1993. The data have been transformed to a fluctuation around zero by subtraction of the mean value 0.2702878 prior to analysis. We have also substituted the gaps of October 1988 and March 1991 with linear interpolation of neighbouring values to achieve a continuous time line. a = 10.4 years, b = 28.5 months, c = 36.6 months and d = 23.3 months.

The spectral analysis of the monthly means of the Mg II index is shown in Figure 28. We have transformed the data to a fluctuation around zero by subtracting the mean value of 0.2702878 before the analysis. The labeled peaks in the figure have corresponding periods of $a = 10.4$ years, $b = 28.5$ months, $c = 36.6$ months and $d = 23.3$ months. The three quasi-biennial peaks may be signs of a single oscillation varying in time as discussed in Section 5.3.4 regarding solar activity.

In Figure 27 the monthly mean Mg II index is plotted together with a 5-month moving average and a 20-month moving average representing the mean trend.

To be able to compare the solar UV fluctuations with what was found for the U/P ratio the two quantities are plotted together in Figure 29. The UV fluctuations appear to contain a superposition of a frequency component with a period that is significantly shorter than two years, and seem to vanish during solar activity minimum. During the

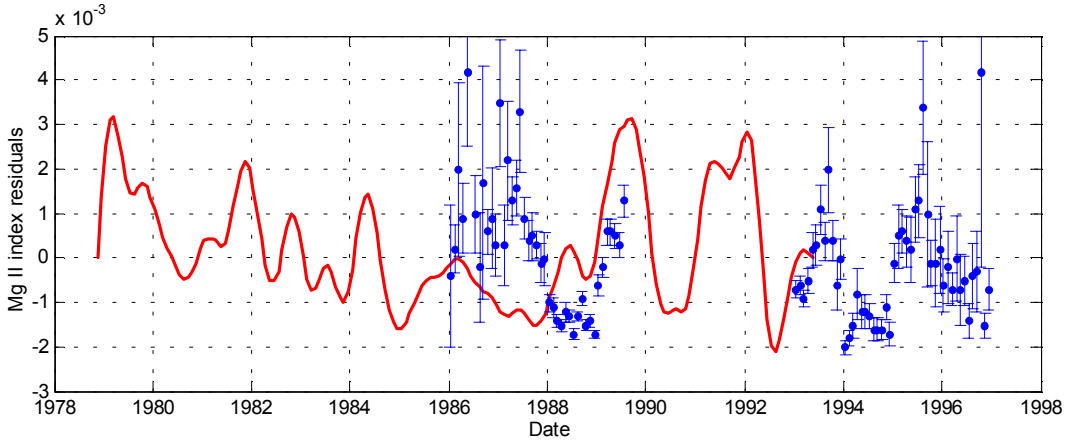

Figure 29: Residuals of the Mg II index 5-month moving average minus the mean trend, plotted as a thick, solid line. The U/P fluctuation is shown in arbitrary scale as monthly means with error-bars corresponding to two standard errors.

overlapping intervals, the U/P fluctuation lags the variations of solar UV irradiance with a couple of months, the number of which is difficult to determine. A more detailed comparison must wait until additional years of sunspot data in the DPD are available.

5.7 Shear Oscillation of the Solar Tachocline

The internal structure of the Sun in the layers below the photosphere is invisible for direct observations. Helioseismology, though, provides a way to indirectly measure internal properties by supplying possible limits for the temperature, chemical composition, and motions. This is done by long-term, very precise, continuous observations of spectral lines in the light emitted at the surface of the Sun. A mid-term periodicity, described in Section 5.7.4, with its strongest spectral peak at a frequency corresponding to a period of about 1.3 years has been reported by Howe et al. [Ho00] in the year 2000.

5.7.1 Principles of Helioseismology

Since the Sun is a sphere of hot gas, its interior transmits sound very well. Convection acts like a random clapper causing millions of distinct sound waves to propagate. Destructive interference damps all but the resonant frequencies and transforms the random convective noise into a very rich line spectrum in the 5-minute range. Thus, resonant sound waves are propagating along different paths throughout the Sun, revealing themselves as Doppler shifts in spectral lines. The periods of the sound waves depend on their propagation speeds, which in turn depend on temperature and chemical composition, and the depths of the resonant cavities. The large number of resonant modes, with different cavities, allows a model reconstruction of different solar parameters with narrow limits for the parameters from just below the surface down to the very core of the Sun. The oscillations are usually called p-modes, p for pressure, since sound waves are propagating pressure perturbations.

Solar internal rotation imparts a clear signature, a splitting, to the oscillation periods. A sound wave propagating in the direction of the rotating gas will have a higher absolute speed than an identical wave propagating against the flow. This splits up the original single oscillation frequency into a measurable pair of close but distinct frequencies, the

wave traveling in the direction of rotation having the higher frequency. Since the photospheric equatorial rotation period is about 25 days and the p-modes have periods of the order five minutes, the splitting is one part in ten thousand [La00].

5.7.2 Helioseismologic Measurements

There are several complementary experiments probing the deep solar interior by nearly continuously recording full-disk Doppler velocity images of the Sun. One of them is the Global Oscillation Network Group (GONG), which consists of six ground-based observatories with each site representing one of six longitudinal bands. These allow the network to make observations of the Sun 24 hours a day during suitable meteorological conditions. Another experiment is a Michelson Doppler Imager (MDI) located on board the SOHO spacecraft at the inner Lagrangian point.

5.7.3 Solar Internal Rotation

Howe et al. [Ho00] have used GONG and MDI data for a 4.5-year time span, May 1995 to November 1999, to determine the radial and latitude dependence of the solar internal rotation from frequency splittings by two different inversion techniques. These methods are called subtractive Optimally Localized Averaging (OLA) and Regularized Least Squares fitting (RLS).

Figure 30 shows the time average of the solar rotation speed as a function of depth at five different latitudes. The plot has been produced by RLS methods used on GONG frequency splittings. Below ∼0.65 R the Sun's rotation rate seems to be independent of latitude whereas further out latitude dependence is strong. In the region between 0.63 R and 0.72 R, usually called the tachocline, the rotation rate is varying with radial distance from the centre at all latitudes. The differences in speed between the adjacent layers give rise to shearing forces on the intervening gas. This region is thought to play an important role in the cyclic generation of solar magnetic fields.

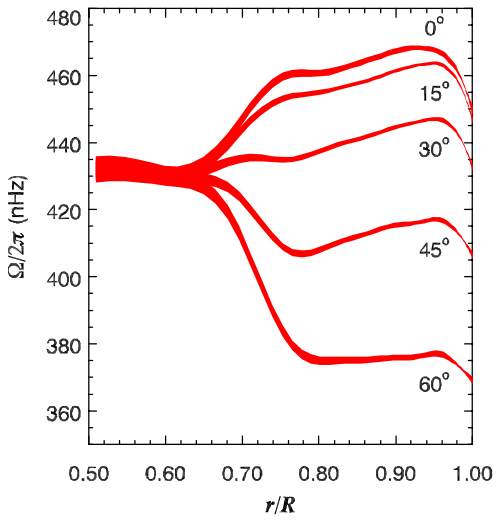

Figure 30: Average rotation rate as a function of radial distance from the solar centre in fractions of the solar radius R, derived from GONG data by RLS methods [Ho00]. The rotation rate is different for different latitudes outside ∼0.65 R. The tachocline region is considered to have its centre at 0.69 R.

5.7.4 Comparisons with the U/P Periodicity

In their investigations of solar internal rotation, Howe et al. have detected a variation of the rotation rate with time near the base of the convective zone [Ho00]. This oscillating behaviour is shown for a latitude of 0° in the lower convective zone at 0.72 R and in the upper radiative interior at 0.63 R, in the upper and lower parts of the right panel of Figure 31. There are variations also at other latitudes, but the signal is most clear at the solar equatorial plane and has a strongest spectral peak at a frequency corresponding to a period of 1.3 years. The oscillation above and below the tachocline region is out of phase. We will therefore use the term shear oscillation in future references to this phenomenon.

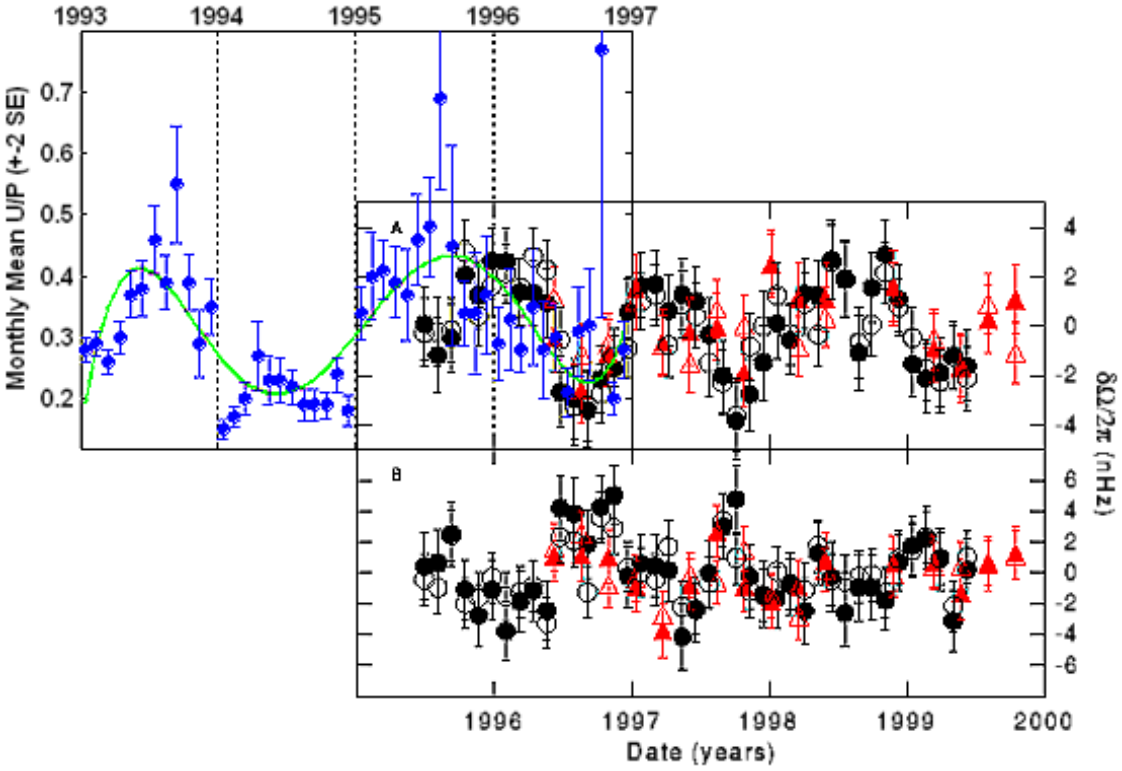

Figure 31: The left panel shows monthly means of the U/P area ratio in blue with error-bars corresponding to two standard errors. The green curve is a polynomial f it described in Section 3.2.4. The right panel shows solar rotation rate residuals derived from GONG-RLS (black filled circles), GONG-OLA (black open circles), MDI-RLS (red filled triangles) and MDI-OLA (red open triangles) from Howe et al. [Ho00]. The upper part contains the residuals at 0.72 R radial distance from the centre and the lower part the residuals found at 0.63 R, both at a solar latitude of 0°.

In the left panel of Figure 31, the U/P ratio fluctuation is shown as monthly means and a polynomial fit described in Section 3.2.4. In the short overlapping interval of the U/P ratio and the shear oscillation, the two periodicities are in phase, but the U/P ratio has a much longer length of cycle. Since the interval of overlap is very short, a connection between the periodicities might anyhow not need to be excluded. The period of 1.3 years found for the shear oscillation by Howe et al. is tentative since the data set analyzed is quite short, 4.5 years, compared to the cycle length. Also, the similar shear oscillations at a solar latitude of, e.g. 30°, have cycle lengths that seem to vary with time more than the equatorial variations shown in Figure 31. However, the equatorial oscillation is varying too: the length of the cycle with maximum during 1998 seems to be longer than

the cycle before. The U/P ratio fluctuations are as described in Section 3.2.4 also showing a varying cycle length, but with an on average longer mean period. When further data for both these periodicities are evaluated any similarities will be possible to examine in more detail. For now we can just conclude that any similarity could be explained by the magnetic flux tubes of sunspots having their origin in the convective zone and the flux tube dynamics maybe being affected by the state of the tachocline region.

5.8 Heliospheric Mid-Term Periodicities

There are several recent investigations made on periodicities in various heliospheric parameters like solar wind speed, interplanetary magnetic field, cosmic ray intensity and geomagnetic activity claiming the existence of mid-term periodicities. These are, however, not new phenomena. A spectral peak in the geomagnetic Ap-index at 1.47 years was reported already in 1972 by Frasier-Smith [Fr72] and what was considered to be an unexpected peak near 1.4 years was found in the Swedish visual aurora records 1983 by Silverman and Shapiro [Si83].

5.8.1 Solar Wind Speed

The solar wind speed as measured at the orbit of Venus, 0.72 AU from the Sun, and the measurements of Voyager 2 and other deep-space probes in the outer heliosphere contain large variations at a period of about 1.3 years, similar to what is found at the Earth's orbit. This suggests the variation not to arise from the merging of structures during solar wind evolution, but rather to be of solar origin as it persists throughout such diverse parts of our solar system. Mursula et al. [Mu02] point out that the similar length of the period compared with what has been found for the shear oscillation of the solar tachocline described in Section 5.7.4, may mean that the heliospheric variations are related to the solar dynamo and the emergence of magnetic flux.

5.8.2 Geomagnetic Activity

Mursula and Zieger have compared solar wind speed data recorded from 1964 onwards at the orbit of the Earth with the Kp-index of geomagnetic activity, available from 1932. Figure 32, published in their paper [Mu00], shows the power spectral density of the two parameters as functions of time, the solar wind speed in the upper panel and the Kp-index in the lower panel. The parameters show similar patterns of spectral density except in the early seventies. During that time the solar wind speed measurements contain long gaps whereas the data series of geomagnetic activity provide more continuous observations and are therefore considered more reliable. During even solar cycles there is a 1.2–1.4-year periodicity, which is strongest in cycle 22. During odd solar cycles the strongest intensities are found for periods of 1.5-1.7 years. The pattern of alternating periodicities in the Kp-index of geomagnetic activity seems to be systematic and implies yet another difference between even and odd solar cycles. Mursula and Zieger draw an interesting conclusion valid if the periodicity is of solar origin and not developed in space. If alternations can be considered systematic also for the solar wind, as the similarities between the parameters when simultaneous measurements are available give a hint of, they might be signs of different development of coronal holes during even and odd cycles. This is due to the coronal holes considered to be the sources of the fast component of the solar wind.

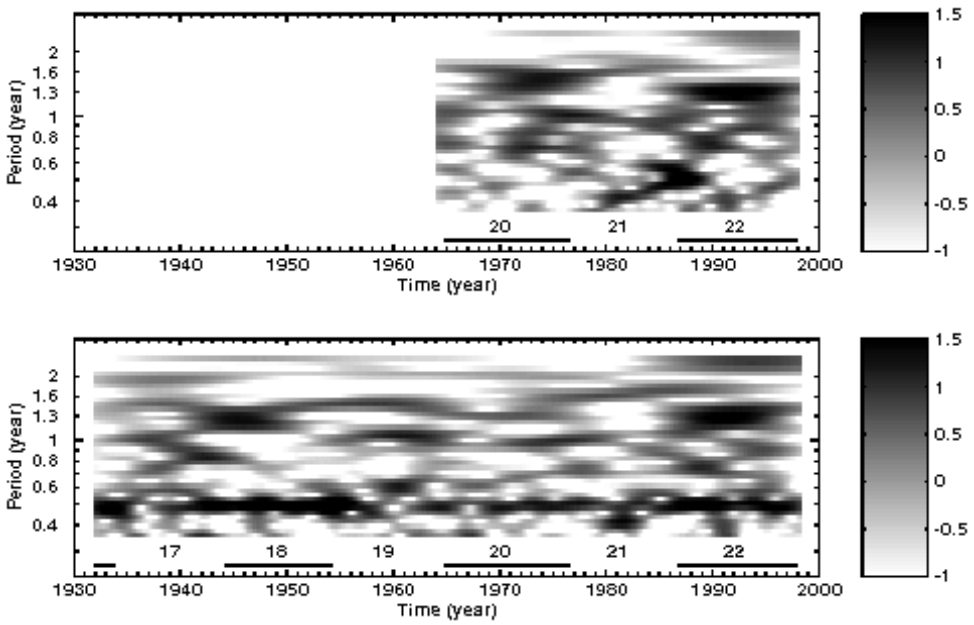

Figure 32: Power spectral density as a function of time for the solar wind speed (upper panel) and the Kp-index of geomagnetic activity (lower panel) from Mursula and Zieger [Mu00].

Mursula et al. have in a later investigation studied the extended aa-index of geomagnetic activity, which is a combination of the aa-index and the Ak(Hel) index derived from early magnetic observations made at Helsinki, Finland [Mu02]. Also in this data set two different spectral density enhancements are found in an alternating pattern. A dominating shorter 1.2-1.3-year periodicity is strongest in the even cycles and a longer 1.5-1.7-year periodicity, usually weaker than the previously mentioned, is dominating during some intervals, particularly odd solar cycles. The spectral power in the 1.2-1.8-year range is furthermore found to be in correlation with the amplitude of the solar cycle as defined by relative sunspot numbers: during low-amplitude cycles there were hardly any power at all for periods within the range.

The long-time behaviour of the extended aa-index is illustrated in Figure 33, from [Mu02]. The time-line of the figure extends from 1844 to 2001. In the upper panel of the figure the monthly relative sunspot number with cycle numbers indicated illustrate the state of the solar activity, whereas the spectral power of the extended aa-index for periods between 0.75 and 2.3 years is shown in the lower panel. Figure 33 will be discussed further in the context of comparisons with the U/P periodicity in Section 5.8.3.

5.8.3 Comparisons with the U/P Periodicity

The established heliospheric mid-term periodicities mentioned previously all have cycle-lengths somewhat shorter than found for the U/P area ratio. The shortest cycle length of the latter is of the same order as the longest described heliospheric periodicity with a period around 1.7 years. However, there is also a quasi-biennial period present in the geomagnetic indices, which Mursula et al. do not consider being of solar origin. This period is seen as high spectral power in a continuous narrow frequency range, with a period very close to 2.0 years, in the lower panel of Figure 33. The periodicity exists during the whole time-span in contrast to the 1.2-1.8 year range for which the power

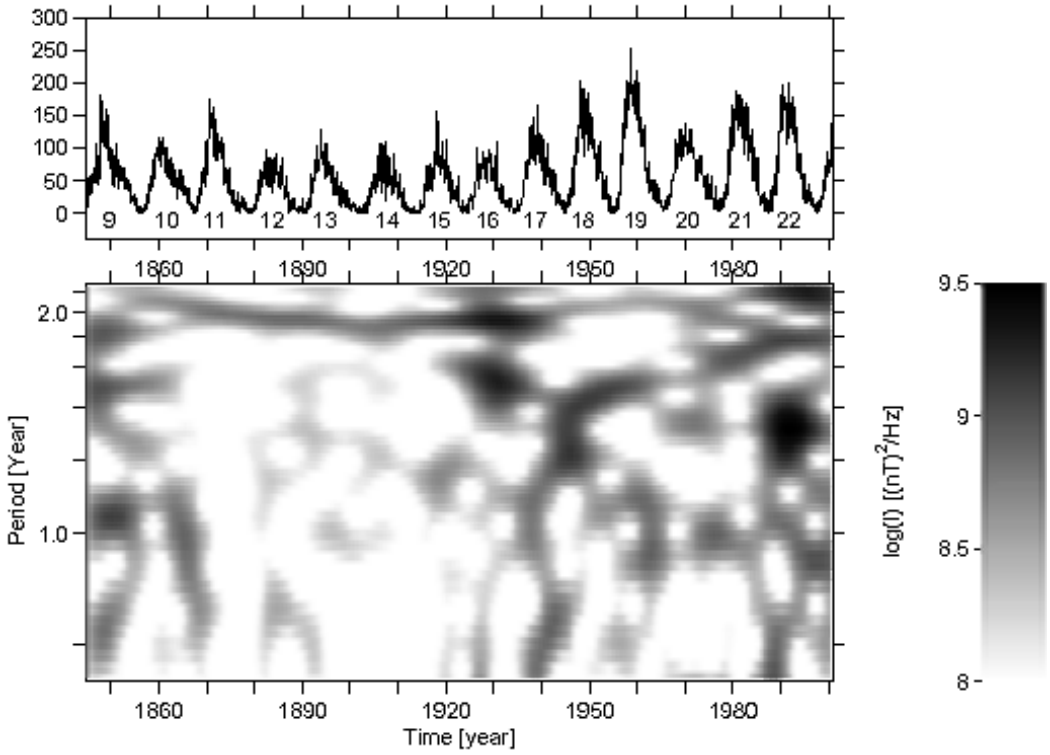

Figure 33: The monthly relative sunspot number, with the numbers of the cycles indicated, illustrating the state of the solar activity (upper panel) and the power spectral density of the extended aa-index of geomagnetic activity as a function of time (lower panel), from Mursula et al. [Mu02].

disappears between the 1870:s and 1920:s.

Also the Kp-index contains a quasi-biennial term seen as a vague power enhancement at 2.0 years through the whole time-span in the lower panel of Figure 32. In the same figure, but in the upper panel showing the dynamic power spectrum of the solar wind speed, no corresponding quasi-biennial signal is found. As the 2.0-year periodicity is not observed in the solar wind speed data, Mursula et al. conclude that the signals with this period in the geomagnetic indices not are of solar origin [Mu02]. Instead they relate them to quasi-biennial oscillations of atmospheric winds

The quasi-biennial periodicity found for the U/P area ratio has hitherto only been possible to detect during minima, ascending and descending phases of a solar activity cycle due to the lack of digitized observations during maxima. If connecting the two available intervals of the DPD with one sinusoidal curve of constant frequency, the most suitable period is slightly longer than two years, as was described in Section 3.2.3. It is therefore interesting to note that a signal with this period consistently exists in the geomagnetic activity measurements. The mechanism of that signal is yet unclear and it is not expected to be of solar origin, but as was described in Section 5.3, quasi-biennial behaviour is in some contexts considered to be the main feature of the fine structure of solar activity.

6 The Quasi-Biennial Oscillation of the Earth Atmosphere

The term 'Quasi-Biennial Oscillation' (QBO) was created in 1964 by Angell and Korshover to describe the phenomenon of reversal of wind directions in the equatorial stratosphere [An64]. For about one year the prevailing wind direction is easterly while it during the following year is westerly. The dynamical time-behaviour of the QBO is seen in Figure 34, which shows a cross-section of the monthly mean equatorial zonal wind with height on the vertical axis and time on the horizontal axis, created by Kapala and Mächel at the Meteorological Institute of the University of Bonn,. Wind speed and direction is colour coded according to the colour bar at the bottom of the figure. The full cycle length of the QBO varies between 20 and 37 months and has a 28.5-month period on average. The phase of the cycle depends on height, low heights lagging the higher regions. Figure 34 has been derived from NCEP/NCAR[4] reanalyses.

According to Hoyt and Schatten [Ho97] the QBO arises strictly from internal interactions between the Pacific Ocean and the atmosphere and no solar forcing is postulated for the phenomenon. However, this is still a matter under investigation, since the mechanisms are not fully understood. We have plotted the U/P area ratio fluctuations from both the DPD and the DPR together with the QBO in Figure 34. When comparing the U/P variations with how the QBO behaves at any arbitrary fixed atmospheric height, it is seen that the two fluctuations alternately are in and out of phase during different intervals. We do not see any physical reason for a connection between the two diverse phenomena.

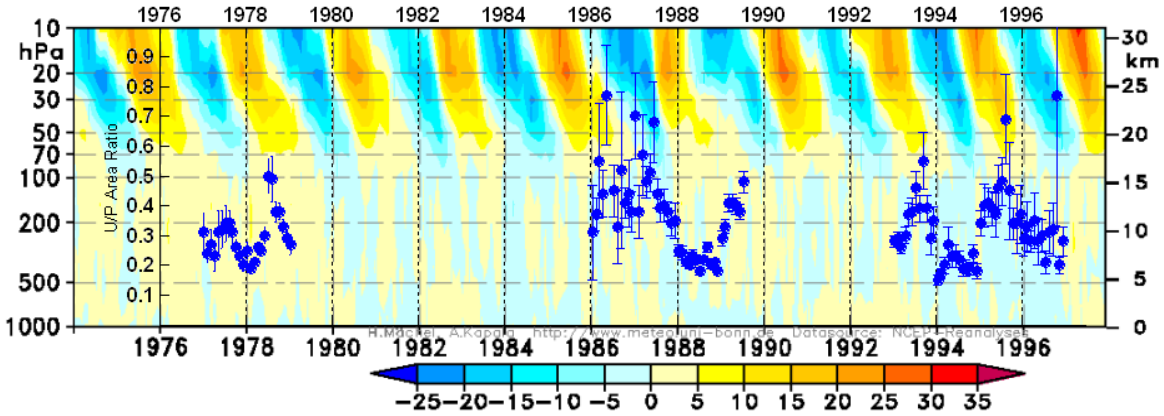

Figure 34: Monthly mean equatorial zonal wind derived from NCEP/NCAR reanalyses by Kapala and Mächel, together with the U/P area ratio from both the DPR and the DPD, presented as monthly means with error-bars corresponding to two standard errors.

[4] National Centers for Environmental Protection/National Center for Atmospheric Research

7 Conclusions

This section will contain only our general remarks on the work at hand. Conclusions and discussions about the results of each investigated subject, as the different quantitative comparisons between the U/P fluctuations and other solar mid-term periodicities, can be found in the subsections named *Our analyses*, *Discussions* or *Comparisons with the U/P Periodicity* in the sections of the subjects.

According to Hoyt and Schatten [Ho97], the importance of a periodicity can be expected to be inversely proportional to the difficulty of discovering it, or the computing labour needed to extract it from a given data set. There is an inverted logic laying in this statement due to the fact that a periodicity can be defined as important by the significance and clarity of its primary or secondary effects, which usually are the reasons for the periodicity being easily discovered in the first place. The statement is logical though, in the sense that other fluctuations of the same phenomenon, which by some reasons are not as easy to detect but may be as important, can only be found when the easiest detectable periodicity is known a priori, since it needs to be taken into account for the other fluctuations to be noticed. Therefore the most easily detectable periodicity, even if it is not intriguing by itself, may be the necessary key to further research in any new field.

Our main point of this discussion is to draw attention to what is distinguishing the U/P area ratio fluctuation from other fluctuations associated with solar activity: the main period of the U/P fluctuation is quasi-biennial and it does not show 11-year cyclic dependence during the interval of the available DPD, which is the most striking feature of solar activity indices. Another merit is its unambiguously clear signature in comparison with even the most seriously considered solar mid-term periodicities, as the shear-oscillation of the solar tachocline. The latter is only really explicit at one certain latitude out of several investigated solar latitudes, the solar equatorial plane, but still regarded a signature of a global solar mechanism. This should tell us something about the significance of the U/P periodicity, if the quasi-biennial period found is not caused by some overlooked, twisted discrepancy in the DPD, which is most unlikely.

We are not able to draw any firm conclusions of the physical interpretation of the U/P fluctuation, neither are we able to point out any specific other solar mid-term periodicity that shows an over-all consistent time-behaviour in comparison with it. The investigations presented in this report may rather be seen as a pre-study for further investigations to come when additional years of sunspot data can be evaluated.

Acknowledgments

The solar radius data used have been produced by Francis Laclare et al. at Observatoire de la Côte d'Azur (CERGA, URA 1360) (France).

The solar irradiance data used are version 19 of a corrected composite irradiance data set received through Weltstrahlungszentrum, Physikalisch-Meteorologisches Observatorium Davos (PMOD/WRC), Switzerland. It partly contains unpublished data from the VIRGO Experiment on board the cooperative ESA/NASA[5] Mission SOHO.

Solar ultra-violet irradiance data have been produced by Richard P. Cebula and Matthew T. DeLand of Hughes STX Corporation. Their data set and the International Relative Sunspot Numbers were acquired through the National Geophysical Data Center of the National Oceanic & Atmospheric Administration (NOAA NGDC), USA.

During our work we have made great use of the NASA Astrophysics Data System Abstract Service (NASA ADS), USA for finding relevant information.

[5] European Space Agency/National Aeronautics and Space Administration

References

[An64] J.K. Angell, J. Korshover; *Quasi-Biennial Variations in Temperature, Total Ozone and Tropopause Height,* Journal of Atmospheric Sciences 21, p. 479-492 (1964)

[Ap85a] E. M. Apostolov; *Quasi-Biennial Oscillation in Sunspot Activity*; Bulletin of the Astronomical Institutes of Czechoslovakia 36, p. 97-102 (1985)

[Ap85b] E. M. Apostolov, V Letfus; *Quasi-Biennial Oscillations of the Green Corona Intensity*; Bulletin of the Astronomical Institutes of Czechoslovakia 36, p. 199-205 (1985)

[Ba00] G. A. Bazilevskaya, M. B. Krainev et al.; *Structure of the Maximum Phase of Solar Cycles 21 and 22*; Solar Physics 197, p. 157-174 (2000)

[Ca02] B. Caccin, V. Penza; *Line-depth and T_{eff} variations with the solar cycle*; in H. Sawaya-Lacoste (editor); *Proceedings of the Second Solar Cycle and Space Weather Euroconference, 24 - 29 September 2001, Vico Equense, Italy*; ESA SP-477, p. 205 - 208, ESA Publications Division, Noordwijk, the Netherlands (2002); ISBN 92-9092-749-6

[Cal02] D. K. Callebaut, V. I. Makarov, A. G. Tlatov; *Gravitational Energy, Solar Radius and Solar Cycle*; in H. Sawaya-Lacoste (editor); *Proceedings of the Second Solar Cycle and Space Weather Euroconference, 24 - 29 September 2001, Vico Equense, Italy*; ESA SP-477, p. 205 - 208, ESA Publications Division, Noordwijk, the Netherlands (2002); ISBN 92-9092-749-6

[Co87] M. Collados, J. C. del Toro Iniesta, M. Vazquez; *A Statistical Study of the Geometrical Wilson Effect*; Solar Physics 112 , p. 281-293 (1987)

[Dj93] D. Djurović, P. Pâquet; *Quasi-Biennial Oscillation in Green Corona Activity and Earth's Rotation;* Astronomy and Astrophysics 277, p. 669-676 (1993)

[Dj96] D. Djurović, P. Pâquet; *The Common Oscillations of Solar Activity, the Geomagnetic Field and the Earth's Rotation;* Solar Physics *167, p. 427-439 (1996)*

[Fo90] P. Foukal; *Solar Astrophysics*; John Wiley & Sons, Inc.; USA (1990); ISBN: 0-471-83935-3

[Fr72] A. C. Fraser-Smith; *Spectrum of the Geomagnetic Activity Index Ap;* Journal of Geophysical Research, Vol. 77, No. 22, p. 4209-4220 (1972)

[Fr97] P. Frick, D. Galyagin, D. V. Hoyt et al. *Wavelet Analysis of Solar Activity Recorded by Sunspot Groups*; Astronomy & Astrophysics 328, p. 670-681 (1997)

[Fr98] C. Frölich, J. Lean: *The Sun's Total Irradiance: Cycles, Trends and Related Climate Change Uncertainties since 1976*; Geophysical Research Letters 25, p. 4399-4380 (1998)

[Gi71] O. Gingerich, R. W. Noyes, W. Kalkofen, Y. Cuny; *The Harvard-Smithsonian Reference Atmosphere;* Solar Physics 18, p. 347-365 (1971)

[Gr97a] D. F. Gray, W. C. Livingston; *Monitoring the Solar Temperature: Spectroscopic Temperature Variations of the Sun*; The Astrophysical Journal 474, p. 802-809 (1997)

[Gr97b] D. F. Gray, W. C. Livingston; *Monitoring the Solar Temperature: Spectroscopic Temperature Variations of the Sun;* The Astrophysical Journal 484, p.511 (1997)

[He86] D. F. Heath, B. M. Schlesinger; *The Mg 280 nm Doublet as a Monitor of Changes in Solar Ultraviolet Irradiance;* Journal of Geophysical Research 91, p. 8672-8682 (1986)

[Ho97] D. V. Hoyt, K. H. Schatten; *The Role of the Sun in Climate Change*; Oxford University Press, USA (1997); ISBN: 0-19-509414-X

[Ho98] D. V. Hoyt, K. H. Schatten; *Group Sunspot Numbers: A New Solar Activity Reconstruction*; Solar Physics 181, p. 491-512 (1998)

[Ho00] R. Howe, J. Christensen-Dalsgaard; *Dynamic Variations at the Base of the Solar Convection Zone*; Science 31 Mars 2000, Vol. 287, No. 5462, p. 2456-2460 (2000)

[Ko02] E. V. Kononovich, S. A. Krasotkin; *Quasi-Biennial Oscillations as the Main Property of the Solar Activity*; in H. Sawaya-Lacoste (editor); *Proceedings of the Second Solar Cycle and Space Weather Euroconference, 24 - 29 September 2001, Vico Equense, Italy*; ESA SP-477, p. 205 - 208, ESA Publications Division, Noordwijk, the Netherlands (2002); ISBN 92-9092-749-6

[La96] F. Laclare, C. Delmas, J. P. Coin, A. Irbah; *Measurements and Variations of the Solar Diameter*; Solar Physics 166, p. 211-229 (1996)

[La00] K. R. Lang; *The Sun from Space*; Springer-Verlag Berlin Heidelberg; Germany (2000); ISBN: 0941-7834

[Le97] J. Lean; *The Sun's variable radiation and its relevance for Earth;* Annual Review of Astronomy and Astrophysics 35, p. 33-67 (1997)

[Li95] L. Liritzis; *Quasi-Periodic Variation in the Solar Neutrino Flux Revisited;* Solar Physics *161, p.29-47 (1995)*

[Mo01] R. Moussaoui, A. Irbah, E. Fossat et al.; *Spectral Analysis of Solar Diameter Measurements Recorded at Calern Observatory Astrolabe During Two Solar Cycles*; Astronomy and Astrophysics 374, p.1100-1107 (2001)

[Mu00] K. Mursula, B. Zieger; *The 1.3-Year Variation in Solar Wind Speed and Geomagnetic Activity;* Advances in Space Research, Vol. 25, Issue 9, p. 1939-1942 (2000)

[Mu02] K. Mursula, B. Zieger, J. H. Vilppola; *Mid-Term Quasi-Periodicities in Geomagnetic Activity During the Last 15 Solar Cycles: Connection to Solar Dynamo Strength*; Solar Physics (in press)

[Pa79] E. N. Parker; *Sunspots and the Physics of Magnetic Flux Tubes. I. The General Nature of the Sunspot*; The Astrophysical Journal 230, p. 905-913 (1979)

[Pr84] E. R. Priest; *Solar Magnetohydrodynamics*; D. Reidel Publishing Company; Dordrecht, Holland (1984); ISBN: 90-277-1374-X, 90-277-1833-4

[Sa79] K. Sakurai; *Quasi-biennial Variation of the Solar Neutrino Flux and Solar Activity*; Nature 278, p. 146-147 (1979)

[Sa81] K. Sakurai; *Quasi-biennial Periodicity in the Solar Neutrino Flux and its Relation to the Solar Structure*; Solar Physics 74, p. 35-41 (1981)

[Si83] S. M. Silverman, R. Shapiro; *Power Spectral Analysis of Auroral Occurrence Frequency;* Journal of Geophysical Research Vol. 88, No. A8, p. 6310-6316 (1983)

[St97] P. A. Sturrock, G. Walther, M. S. Wheatland; *Search for Periodicities in the Hemestake Solar Neutrino Data*; The Astrophysical Journal 491, p. 409-413 (1997)

[Sz03] C. Szasz; *Fluctuations of the Sunspot Umbra-Penumbra Ratio;* Master's thesis at Luleå University of Technology, Sweden (in press)

[Ul95] R. K. Ulrich and L. Bertello; *Solar-Cycle Dependence of the Sun's Apparent Radius in the Neutral Iron Spectral Line at 525 nm*; Nature 377, p.214-215 (1995)

[Zi88] H. Zirin; *Astrophysics of the Sun*; Cambridge University Press (1988); ISBN: 0-521-316073