# Introduction to the Solar Activity Cycle: Overview of Causes and Consequences

A. Balogh · H.S. Hudson · K. Petrovay · R. von Steiger

Published online: 5 December 2014 © Springer Science+Business Media Dordrecht 2014

Abstract The 11-year activity cycle is a dominant characteristic of the Sun. It is the result of the evolution in time the solar dynamo that generates the solar magnetic field. The nearly periodic variation in the sunspot number has been known since the mid-1800s; as the observations of the Sun broadened to cover an increasing number of phenomena, the same 11-year periodicity was noted in most of them. The discovery of solar magnetic fields introduced a 22-year periodicity, as the magnetic polarities of the polar regions change sign every 11 years. Correlations have been identified and quantified among all the measured parameters, but in most cases such correlations remain empirical rather than grounded in physical processes. This introductory paper and the reviews in the volume describe and discuss the current state of understanding of the causal chains that lead from the variable nature of the solar magnetic fields to the variability of solar phenomena. The solar activity cycle is poorly understood: predictions made for the current Cycle 24 have proved to be generally wrong. However, the re-evaluation of the relationships in the light of unexpected shortcomings is likely to lead to a better physical understanding of solar physics. This will help in the systematic reassessment of solar activity indices and their usefulness in describing and predicting the solar activity cycle.

Keywords Sun · Solar activity · Solar cycle · Solar indices · Solar magnetic field

A. Balogh (🖂)

Imperial College London, London SW7 2AZ, UK e-mail: a.balogh@imperial.ac.uk

H.S. Hudson

Space Sciences Laboratory, University of California, 7 Gauss Way, Berkeley, CA 94720-7450, USA e-mail: hhudson@ssl.berkeley.edu

K. Petrovay Department of Astronomy, Eötvös Loránd University, Pf. 32., 1518 Budapest, Hungary e-mail: K.Petrovay@astro.elte.hu

R. von Steiger International Space Science Institute (ISSI), Hallerstrasse 6, 3012 Bern, Switzerland e-mail: vsteiger@issibern.ch

## 1 Introduction: Correlations and Causal Relations

The solar activity cycle, a near-periodic variation generally close to a period of 11 years, appears prominently in most parameters that describe solar phenomena. Records of variations, originally based on sunspot counts, show that this has been the case for historic times, even if during certain intervals lasting decades and longer sunspots were rare. This demonstrates the existence of a timing engine in the Sun that appears to control or influence all aspects of solar phenomena and also beyond, including phenomena elsewhere in the solar system that are affected by the Sun. In recent decades, understanding the solar activity cycle has progressed from being of purely scientific interest to a practical need because of the recognition that solar phenomena affect the Earth's space environment in ways that can be potentially damaging. It is generally agreed that the timing engine in the Sun causes a nearly periodic generation and evolution of magnetic fields in the solar interior and on the solar surface, more explicitly, that a dynamo mechanism generates the Sun's magnetic field. In the past half century, great improvements in ground-based solar observations have been combined with observations from space-together, these advances have provided an almost unlimited data volume describing solar variability, on all time scales and at all spatial, spectral resolution.

The nature of the solar dynamo remains difficult to determine. But for solar activity, Leighton (1969) suggested that "...a comprehensive test of any model of the solar cycle will ultimately require the knowledge of the strength, spatial distribution, and history of the solar magnetic field spanning a considerable time period. The existing indices of solar activity, while forming a valuable continuous record covering several complete cycles, will probably be of diminishing utility unless they can be reliably related to physically significant quantities". The very broad range of indices now available, following the great increase in the quality and quantity of solar observations, remains subject to this assessment. Although progress in recent decades has been significant in delineating possible interdependences of solar activity indicators as well as the physical basis of such interdependences, in most instances the relationships remain on the phenomenological level and do not reach, generally speaking, a strict physical causal level of rigour.

The best test of causal understanding between related phenomena is the ability to make predictions that can then be verified by observations. The understanding of the solar activity cycle has largely failed this test by not being able to predict the dominant parameters (in particular sunspot numbers) for the current Cycle 24 that followed the already unusual Cycle 23, in particular the long-duration and deep minimum in activity levels between these cycles. See Pesnell (2012) for an overview of the methodology of the predictions used for forecasting the maximum sunspot number in Cycle 24. The summary of the numerous predictions for the maximum of the current Cycle 24 is shown in Fig. 1 (based on Pesnell 2012), together with the sunspot numbers actually observed in Cycle 24 to late 2014. It is generally agreed that Cycle 24 has a smaller maximum than previous recent cycles. Although maximum activity in Cycle 24 has probably, if not certainly, been reached by late 2014, the maximum sunspot number in the cycle remains to be agreed. In any case, Cycle 24 has already proved to be quite different from recent cycles. For a detailed review of the methodologies used in the predictions and their phenomenological and/or physical basis, see Hathaway (2010), Petrovay (2010) and Pesnell (2012).

The striking feature of many of the prediction techniques is their grounding in phenomenology, rather than in Leighton's "physically significant quantities". Given data sets that describe the solar cycle in different phenomena over as many cycles as possible, these methods seek correlations that can be seen to hold over a number of cycles. These correlations are then used as predictors, usually to describe the coming cycle in terms of the



**Fig. 1** In the left panel, the predictions made for the solar activity Cycle 24 are shown, using different methodologies (from Pesnell 2012). The methodologies are listed on the figure; the numbers of predictions that were made using the given methodologies are also given. The red symbols represent the mean value of the predictions, and the vertical bars represent the range of values that were predicted by the different methodologies. In the right panel, the actual values of the monthly sunspot numbers are plotted (in red), together the daily sunspot numbers for 2014 (in blue) indicating the very large scatter that will be smoothed out as the latest data are processed. The green lines represent the latest predicted sunspot numbers for cycle 24 (from http://solarscience.msfc.nasa.gov/images/ssn\_predict\_l.gif). The dashed black lines in the two plots correspond to the largest and smallest predicted maximum in Cycle 24

expected maximum value of the sunspot number and its timing. Predictions for Cycle 24 have uncovered the many potential limitations of the algorithms used and the correlations on which they have been based. It is expected that the deviations from the expected extrapolations will lead to a reassessment of the way in which the physically (at least directly) unrelated phenomena are used in empirical analyses to interpret and to predict solar activity variations.

It is important to review and assess the status of causal understanding of solar activity cycle indicators. This is the task that was undertaken in a Workshop organised by the International Space Science Institute (ISSI), Bern, in November 2013. The resulting reviews (this Issue of Space Science Reviews) provide a comprehensive overview of all aspects of the solar activity cycle and the indices that are used to describe it. A brief introduction to the issues, their historic context and key references are given in this summary paper.

Finally, we restate Leighton's remark in the modern context. We have many well-defined indices that reflect solar magnetic activity in various ways. They are all proxies of some property of the solar magnetic field, but what property is that? The papers in this volume all deal implicitly with this question. We are far from getting answers, but it seems clear that further progress will only come from a self-consistent and physically based combination of improved data and cleverer theory.

#### 2 The Schwabe and Hale Solar Activity Cycles

The Sun may be only one star among the quasi-uncountable number of stars in the Universe, but it is of unique, immediate and vital interest to all life on Earth. Among stars, it has unexceptional properties, but its study in the scientific era has been a central concern in astronomy—the observational and descriptive aspects of the studies, and also in astrophysics—the research into the physical description of its observed properties. One of

the many remarkable observational properties of the Sun is its variability, noted from the earliest telescopic observations four centuries ago. Variability was seen, almost exclusively, in the sunspots—their sizes, shapes, numbers, as the only departure from the "perfection" that had been hypothesised (indeed becoming often an act of faith) in earlier centuries. It is now known that all the properties of the Sun vary in time, with several chains of complex causality suggested in the past century and a half.

Surprisingly, it was only from the mid-19th century that the periodic nature of the variability was recognised as a key feature (for a historical review, see Cliver 2014, this issue). The principal measure, the sunspot number index, was found to vary with a period of about 11 years. Reconstructing past observations and the systematic monitoring of sunspots since that time has provided a time series that, even to date, remains the basic data set on which rests the science of solar variability and much of solar physics (Clette et al. 2014, this issue). Of course, the data set itself, given its importance, has been and remains still now the subject of scrutiny, reworking and even disagreements, but the underlying near-periodicity in this index conditions the way all aspects of solar variability are described and studied. It was the gradually increasing diversity of observational data sets that described other solar phenomena (and even their terrestrial effects), all in greater or smaller ways related to the variations in sunspot numbers, that has led to the concept of solar activity. This concept has become very complex, with large numbers of observables beyond sunspot numbers that can be used to describe the aspects of activity. Progress in solar physics and the science of the Sun's variability and activity have come, over the past century and a half, first from the range, volume and quality of solar observations and the general increase in scientific effort devoted to the Sun, to the interpretation of the observations as well as from the investigations of the underlying physical processes.

Measures of solar activity can be based on many observable and measurable phenomena. All of them have a common characteristic: they are all varying (at least nearly) in phase or anti-phase with the sunspot numbers. Generally speaking, all the observed and measured quantities can be taken to be proxies of some key aspect of the physical causes of variability in solar phenomena. Equally generally, the physical relationships between these measurable parameters are rarely easy to establish—in particular, it is found that the relationships are usually neither simple nor linear (Ermolli et al. 2014, this issue).

The 11-year solar activity cycle, based on the quasi-periodic variations of the sunspot numbers, has been named, after its discoverer, the Schwabe activity cycle (Schwabe 1844). Sunspot cycles are shown in Fig. 2, numbered according to convention, with Cycle 1 starting in 1755. Cycles are counted from one minimum in sunspot numbers to the next minimum. Early in the last century, the discovery and observation of the magnetic fields on the Sun, starting with the interpretation of spectral shifts as caused by the Zeeman effect due to intense magnetic fields in sunspots (Hale 1908), have led to the identification of a cycle with twice the duration of the Schwabe cycle, the solar magnetic, or Hale cycle (Babcock 1959). This 22-year cycle, manifested in magnetic phenomena including the reversal of the poles of the solar magnetic dipole, underlines the fundamental role played by solar magnetism in controlling the solar activity cycle (Petrie et al. 2014, this issue; Wang 2014, this issue; Cliver 2014, this issue). Figure 2 illustrates the two Hale cycles for the recent decades. For the Hale cycle, the magnetic field intensity observed in the polar regions is shown, indicating a reversal of the magnetic polarity of the Sun at about the time in the Schwabe cycle when the number of sunspots has a maximum.

While the global measures of solar activity, primarily the sunspot number, indicate the quasi-periodicity of the underlying mechanism (the solar dynamo, in effect), there are intriguing differences between the two solar hemispheres in detail. Activity is generally not



**Fig. 2** In panel **a** the sunspot number is shown as the most commonly used indicator of solar activity. In panel **b**, the intensity of the galactic cosmic rays measured at Earth is shown, illustrating the dominant 11-year modulation with the Schwabe cycle, and indicating the different modulation patterns in alternate cycles, in phase with the 22-year Hale cycle. In panel **c**, the strength of the solar polar magnetic fields is shown. The polarities reverse every 11 years, producing the 22-year Hale cycle. Note the much reduced strength of the polar fields in the minimum activity phase between Cycles 23 and 24

equal between North and South: the sunspot numbers differ and the two hemispheres show slightly different evolution. Generally speaking, the reversal of the signs of the magnetic poles is not simultaneous, but can differ by many months. The asymmetry and asynchronicity between the hemispheres has been noted to extend to the solar atmosphere and into the heliosphere (Virtanen and Mursula 2014). The question of coupling between the solar hemispheres and its relationship with the evolution and solar activity, together with the wealth of observations that relate to the differences and similarities in activity indices is quite central to the understanding of the whole solar activity process (see Norton et al. 2014, this issue, and references therein).

The solar magnetic field extends into the heliosphere (see, e.g. Balogh and Erdős 2013) where it is entrained by the solar wind, resulting in complex evolutionary structures following the dynamic evolution of the solar wind. A particularly clear link between the solar origin of the magnetic field and the large-scale structures of the heliospheric magnetic field is the interplanetary sector structure. In interplanetary space, a global current sheet (the Heliospheric Current Sheet, or HCS, see Smith 2001), the sector boundary, separates the two large-scale magnetic polarities present in the heliosphere. This structure also has clear signatures at the level of the photosphere, and in the corona in between. The sector structure and the topology of the HCS are related to the evolution, over the Hale cycle, of the solar magnetic field and are important indicators of the state of the heliosphere as a function of solar activity (Hudson et al. 2014, this issue).

An interesting consequence of the extension of solar activity into the heliosphere is the periodic modulation of the flux intensities of high energy cosmic rays, as illustrated in Fig. 2. Galactic cosmic rays entering the heliosphere encounter propagation conditions in the interplanetary medium that vary with the solar activity cycle (see Potgieter 2013; Bazilevskaya

et al. 2014a, this issue). Simply stated, the propagation from the boundaries to the inner heliosphere where the Earth is located involves a greater decrease in cosmic ray intensity at sunspot maximum than at sunspot minimum. Hence the cosmic ray intensity varies in antiphase with the sunspot number, in other words, with solar activity, as can be clearly seen in Fig. 2. However, as can be observed also in Fig. 2, the appearance of modulation cycles of cosmic ray intensities change in alternate 11-year cycles, with a peaked intensity profile followed by a flat-topped intensity profile. This, together with the satisfactory explanation of the observations linking them to the dominant dipole polarity of the sun, is evidence for a 22-year Hale cycle modulation effect.

Cosmic ray intensities are modulated by diffusion and convection effects in interplanetary medium; in effect irregularities in the interplanetary magnetic field and propagating structures lead, statistically, to the diffusion of the high energy particles. The properties of the interplanetary medium change due to solar activity, hence the 11-year variation in intensity. The explanation for the 22-year Hale cycle modulation is to be found in gradient and curvature drift effects on the cosmic ray particle trajectories of the large-scale heliospheric magnetic field. The effects depend on the direction of the magnetic field in the Parker spiral. When the northern heliospheric magnetic field is outward (the polar field is positive, denoted by A > 0, positively charged particles drift inward from the polar regions of the heliosphere to the equatorial plane in the inner heliosphere and the equatorial propagation conditions influence to a lesser extent the access of the particles. When the northern field is directed inward (the polar field is negative, represented by A < 0), positively charged particles enter the inner heliosphere by drifting towards the equatorial regions of the heliosphere, where propagation conditions are controlled by the properties of structures in the heliospheric magnetic field that are co-rotating with the sun. (The constant A is a multiplier in the heliospheric magnetic field model used for cosmic ray modulation studies, see Jokipii et al. 1977.) These effects have been largely confirmed by extensive modelling of the three-dimensional propagation of cosmic rays (Jokipii and Thomas 1981; Kota and Jokipii 1983; Webber et al. 1990).

Although this effect of the 22-year magnetic cycle of solar activity on cosmic rays is well understood, it is interesting to note that a now almost forgotten 22-year recurrent effect had been observed much earlier in the diurnal anisotropy of very high energy cosmic rays (Thambyahpillai and Elliot 1953; Forbush 1969). This effect was also found to be related to the gradients of cosmic rays in the inner heliosphere, and related in turn to the orientation of the sun's magnetic polarity at the activity cycle minima (McCracken et al. 2004).

Generally speaking, many of the intriguing problems in solar physics are related to the causal connections between the Schwabe and Hale cycles. Most of the numerical indicators of solar activity are associated with the 11-year Schwabe cycle, while the Hale cycle seems to reflect the quasi-periodic recurrence of generally complex magnetic phenomena. However, as can be seen in Fig. 3, even the most commonly used and most direct measures of solar activity present a great variability superimposed on the 11- and 22-year periodicity; this is the case for all the measured or derived parameters describing different aspects of solar activity (Cliver 2014, this issue; Ermolli et al. 2014, this issue). Over historical timescales, the Sun's activity has been variable over many timescales, with longer intervals of recurrence (see, e.g. Usoskin 2013), although the 11-year periodicity appears to have survived through all other modulation effects of solar activity. The duration of the cycles is itself quite variable. The average period is close to 11 years, but with a broad distribution around the mean. The cause of such variability in the period remains unclear, but is clearly related to the operation of the solar dynamo which arguably behaves chaotically (Arlt and Weiss 2014, this issue). Equally, the variability in the "strength" of the cycles, measured either by



**Fig. 3** Panel **a** shows the sunspot number, panel **b** the monthly mean intensity of the 10.7 cm radio flux, panel **c** the Total Solar Irradiance, panels **d** and **e** the monthly rates of C-class and X-class solar flares, respectively. The four quantities correspond to very different physical processes on the Sun, with no simple causal connections among them, but all four show the same 11-year periodicity, even though there are differences in the short term variabilities (The sunspot numbers are from the WDC-SILSO, Royal Observatory of Belgium, Brussels; the solar radio flux monthly averages are from Natural Resources Canada, http://www.spaceweather.ca/solarflux/sx-eng.php; the TSI figure shows data from Physikalisch-Metorologisches Observatorium Davos (PMOD); the solar flare rates are adapted from Fig. 3 in Aschwanden and Freeland 2012.)

the number of sunspots at maximum, or by any other activity measure, is unclear. These uncertainties and a lack of understanding of the causal chains of solar phenomena make the prediction of activity indices quite a challenge, as already discussed. Another approach to assess the meaning of the quasi-periodicity is to consider the low order modes of the solar dynamo and use observational evidence (both from sunspot numbers and from the insights about the solar interior provided by helioseismology, see Sect. 4) and explore the nature of the oscillator that matches the observations (Lopes et al. 2014, this issue). A better understanding of a first order dynamo model could lead even to an improvement in predictions, based on the accurate determination of the sunspot numbers, as this is becoming available (Clette et al. 2014, this issue).

A well-established shorter timescale variability in solar activity, the so-called quasibiennial oscillations (QBOs) with the time scale of 0.6–4 years can be clearly seen in such basic indices as sunspot numbers and sunspot areas. Although basically unexplained, this phenomenon has allowed very detailed tests of the operation of the solar dynamo (see Bazilevskaya et al. 2014b, this issue, and references therein). A particularly attractive and potentially fruitful concept is that the QBOs are the effect of the interaction of the dipole and quadrupole terms of the solar dynamo.

## **3** A Selection of Solar Activity Indicators

The diversity of the observable parameters that can be linked to solar activity is very large, and they are all related at least in the sense that their variations follow the nearly 11-year periodicity, and therefore they are all cross-correlated to a greater or smaller extent. However, the correlations do not imply causal relationships, but are only indicative of the underlying driver of the variability that is the solar magnetic field. In Fig. 3, five measured parameters

are shown, all five can be seen to vary in (near) synchronism at the rhythm of the activity cycle. The parameters have been chosen to illustrate that the variations in the parameters are correlated, despite the lack of known physical relationships between them. The parameters are (a) the sunspot number, (b) the integrated solar radio noise at 2800 MHz (the 10.7 cm radiation), (c) the Total Solar Irradiance, (d) monthly rates of C-class solar flares and (e) monthly rates of X-class solar flares.

The historical importance of sunspot counts as a measure of solar variability and the solar activity cycle has already been described. Even if calibrated, it can hardly be said that the Sunspot Number (SSN) is understood physically. In its modern state, it is relatively precise and reproducible. But the physical phenomenon of the sunspot itself has no simple physical theory that explains its nature, for example in the diminution of convective flux that results in its darkness (Archontis 2012; Rempel 2012). Smaller-scale but comparably intense magnetic fields appear in faculae, where the consequences on the emergent intensity differ radically. This juxtaposition of a careful set of observations with an ill-understood quantity, which we term solar magnetic flux emergence through the solar cycle is reviewed by Schmieder et al. (2014, this issue). Given that magnetic fields on the Sun were first discovered in sunspots as already mentioned (Hale 1908), the magnetic fields of sunspots remains a topic of interest as a key observable that is quite intimately related to solar activity (Borrero and Ichimoto 2011). As discussed below, the recent decrease in the average magnetic field of sunspots is a significant indicator of more profound changes in solar activity (Livingston et al. 2012).

The solar radio flux, at a wavelength of 10.7 cm (2.8 GHz), has been measured continuously since 1947 and has been widely used as an indicator (often called F10.7) of solar activity (see, for a recent review, Tapping 2013, and in the context of other solar activity indices, Ermolli et al. 2014, this issue). The match between the SSN and the 10.7 cm radio flux had been very close, once the correct non-linear relationship between the two has been established (see, e.g., Tapping 2013). The 10.7 cm radiation comes from regions of strong magnetic fields on the solar surface and from strong field regions in the corona. These sources provide the variation in synchronism with the activity cycle. Additionally, there is a non-zero flux even at solar minimum, mostly from thermal free-free emission. The close relationship between sunspot numbers and 10.7 cm radio flux, however, has appeared to change over the last complete solar cycle, Cycle 23. The cause of this, briefly discussed below in this Section, may provide important clues to the causal connection between these two indices.

The Total Solar Irradiance (TSI) is a measure of the solar output; it is of immediate relevance as a fundamental parameter in the Earth's energy balance. TSI is the amount of solar electromagnetic radiation per unit area, integrated over all wavelengths, incident on the Earth (measured as  $W m^{-2}$ , and normalised at 1 AU). It has been measured by space instruments for the past almost 40 years, but due to the difficulty of making calibrated absolute measurements, the construction of a generally accepted composite time series of the observations (using data from the different instruments not overlapping in time) has only been possible in the recent past (see, e.g. Frölich 2012; Ball et al. 2012). TSI varies on many timescales, from minutes to years, in particular in phase with the sunspot numbers. Figure 3, panel *c* shows the daily variations in TSI, together with a smoothed average; the amplitude of the daily variability is clearly much greater than the long-term variability. Most of its variability correlates well with sunspot and facular areas. However, the very complex variations in the spectrum of the solar irradiation as a function of the activity cycle, in particular as a function of the strength, distribution and short- and medium-term variability of the magnetic features remain to be explained (Yeo et al. 2014, this issue). Progress will be made by more

accurate measurements, in particular spectral measurements, combined with more detailed models of the spectral irradiance.

Solar flares are short duration energetic outbursts; here short duration means minutes to tens of minutes, when magnetic energy stored in topologically complex magnetic fields above the photosphere is suddenly released. Solar flares produce energetic particles, plasma heating and a broad spectrum of electromagnetic radiation. Solar flare observations have provided considerable quantities of data that have increased in volume and complexity as new observing facilities have become operational (Fletcher et al. 2011). The occurrence rate of solar flares is linked to solar activity as measured by, for instance, the sunspot number, but in fact there are differences in terms of the phase of the sunspot cycle, as well as its intensity. Although, as shown in Fig. 3, panel d, the rates of C-class solar flares (relatively small flares, without noticeable terrestrial effects) varies with an approximate periodicity of 11 years, the shape of the frequency distribution differs significantly from other measures of solar activity. The relatively small number of the largest, X-class flares, shown in Fig. 3, panel e, have an occurrence frequency that only indicates that such flares occur away from solar minimum activity, but not matching the profile of other activity indicators, such as the sunspot numbers. (C- and X-class flare occurrence frequency in Fig. 3 is based on Fig. 3 in Aschwanden and Freeland 2012.)

This brief and illustrative discussion of apparently very different phenomena and processes that show the periodic solar activity variations indicates strongly that the underlying mechanism driving solar activity is an all-pervasive process that affects all the characteristic observable solar phenomena. Even though the physical causal relationship among the indicators is largely unknown, and in fact in many cases may not even exist, there are, undoubtedly, some dominant causal chains that may act in parallel. This means that the observed phenomena are driven by the yet to be established physical processes, some of which operate independently, but eventually all of them originate in a single, central driving mechanism which can only be the generation mechanism of the Sun's magnetic field.

An example of the way in which future progress can be made concerns the study of the deviation of the empirical function linking the sunspot number and the 10.7 cm solar radio noise, F10.7, during the last Cycle 23 (Svalgaard and Hudson 2010). The deviation is indicated in Fig. 4(left panel). As the F10.7 data and the sunspot number had been considered to be a good proxy for each other and for quantifying solar activity, the deviation observed shows that the physical link between "solar activity" and sunspot numbers on the one hand, and the F10.7 flux on the other is neither direct nor single-valued. Given the dependence of the of the radio noise on total magnetic field, the result has been linked to the decrease in the magnetic field of sunspots (Penn and Livingston 2006; Svalgaard and Hudson 2010; Pevtsov et al. 2011; Livingston et al. 2012). The sunspot formation fraction (SFF), a parameter introduced by Livingston et al. (2012) is the ratio of the sunspot number to that predicted using the previously established relationship to the F10.7 flux. As can be seen from Fig. 4, the SFF remained close to 1 (indicating a high degree of predictability) until about 1995. After that time, the sunspot number decreases relative to the prediction; this change occurs coincidentally with qualitative and quantitative changes in other solar activity indicators. The close relationship observed previously, from 1947 when the F10.7 data series starts, covered almost five Schwabe cycles and more than two Hale magnetic activity cycles and could be considered to be well established, even if not completely understood. The significant deviation through the last cycle from the previous behaviour can be used to investigate the underlying parametric causes of it and hence to clarify better the causal links to the extent that they can be identified. Livingston et al. (2012) found a systematic decrease in the magnetic field strength of sunspot umbrae since about 1998. A consequence of this is



**Fig. 4** *In the left panel*, the solar 10.7 cm radio flux is plotted vs. sunspot numbers. The radio flux has been used recently to indicate the level of solar activity, partly because of its close numerical match to the sunspot number variations (once normalized) and partly because it is closely linked to the solar effects on the Earth's upper atmosphere. However, the relationship between the two quantities is non-linear and appears to have changed significantly during the last activity cycle (based on Fig. 12 in Svalgaard 2009). *In the right panel*, the possible cause of the change is shown, based on a new parameter, the sunspot formation fraction introduced by Livingston et al. (2012)

the decrease in the detectability of weak-field sunspots that correlates well with the decrease in the SFF parameter over the same epoch.

#### 4 The Source of Variability: The Solar Magnetic Field

The engine of solar variability is undoubtedly the solar magnetic field. However, the range of magnetic structures effectively covers all temporal and spatial scales, from still unresolved photospheric structures to the global field which, if not simple, nevertheless provides evidence for global organisation. The dominant quasi-periodic variations represented by both Schwabe and Hale activity cycles appear to be strong evidence for global structuring and evolution. A solar dynamo is essential for understanding these determining features of the phenomenology of solar magnetism and solar activity. However, despite steady progress observationally, theoretically and, most importantly, in realistic numerical simulations, there are fundamental areas of ignorance concerning the generation and evolution of magnetic fields. In the recent past, the paradigm of Flux Transport Dynamos (FTD) has dominated research efforts. Originally proposed by Wang and Sheeley (1991) to build on the generally well-regarded model by Babcock and Leighton (for a recent assessment, see Charbonneau 2007), it has provided a framework in which the kinematics and dynamics of magnetic field generation have been successfully investigated both observationally and in modelling terms (Karak et al. 2014, this issue). A link between the magnetic field dynamics in the solar interior and in the photosphere and corona is provided by the concept of magnetic helicity (Pevtsov et al. 2014, this issue). This concept has been fruitfully used to explain such activity related phenomena as Joy's Law on the directional orientation of sunspot pairs of opposite polarities and the orientation of solar active regions that evolve on timescales much shorter than the activity cycle.

Nevertheless, even if much of the theoretical basis and many of the numerical tools for modelling the solar dynamo have been developed, the problem remains very complex (see, e.g., Charbonneau 2010). The outer third or so of the Sun is convective, involving the transport of solar material towards the surface in an ill-understood manner. The relevant physical parameters of the solar material are not known in sufficient detail, and the kinetics cannot be as yet deduced or modelled to the extent that the magnetic field generation process can be unambiguously deduced. The key to progress is a better understanding of the Sun's convection zone, as well as the shear layers at the bottom of the convection zone and close to the solar surface. Sounding the material properties inside the solar dynamo. The study of the differences between the solar hemispheres and the way the development of activity is coupled between North and South, as already mentioned (Norton et al. 2014, this issue), highlights the complexity of the physical modelling the dynamo, while it also provides observational tools and even constraints on the models.

For the past 40 years, the use and continued improvement of a fundamentally important tool, helioseismology, has proved exceptionally powerful for investigating the interior of the Sun (see Broomhall et al. 2014, this issue). Helioseismology, both ground- and space-based, has provided the data on the internal structure, dynamics and physical properties of the Sun that have opened up the possibility to link structures and processes in its interior to its visible properties and its activity cycle (Thompson 2010; Komm 2010).

Given the observations of the changes in activity between the earlier cycles in the space age and in the declining phase of Cycle 23 leading to the current Cycle 24, it is possible to explore corresponding changes in a range of previously well-established relationships between solar activity indicators, as already shown in Sect. 3 with reference to the sunspot number and the solar 10.7 cm radio noise intensity. These studies will help in identifying actual physical processes behind the empirical relationships and thus better to distinguish between causes and consequences in solar activity phenomena.

The fundamental importance of the solar dynamo mechanism in the generation of all the activity phenomena and, in turn, of the role of the dynamics and physical characteristics of the solar interior, helioseimologic results concerning the changes through the solar cycle are primary indicators of the underlying causes of changes in activity. Shifts in the frequencies of solar oscillations track the 10.7 cm radio noise quite closely, see Fig. 5(a). The frequency shifts have been compared in fact to a range of solar activity indicators by Chaplin et al. (2007). Their conclusion, based on data sets that ended as the minimum activity phase between Cycles 23 and 24 just started, was that the frequency shifts were best correlated with activity indices that are sensitive to both weak and strong surface magnetic fields, e.g. the 10.7 cm radio flux, rather than those which are only sensitive to strong fields, e.g. the sunspot number. The good match is shown in Fig. 5(a) (after Broomhall et al. 2009). This implies that the frequency shifts correspond to magnetic field changes globally, rather than in the high-activity belt in the Sun's equatorial regions. Basu et al. (2012) have investigated the dependence of activity-related frequency shifts as a function of the low-, mid- and highfrequency modes of solar oscillations. Their results, shown in Fig. 5(b) indicate strongly that some changes in solar structure and the magnetic field were already taking place during Cycle 23, when other indicators had not yet clearly foreseen the long minimum between Cycles 23 and 24 and the very low activity level of Cycle 24. Their interpretation of these results is (also exposed in detail by Broomhall et al. 2014, this issue) that the Sun's subsurface





magnetic layer (which is thought to be where structural differences and magnetic fields lead to the manifestations of solar activity) thinned, starting from Cycle 22. The implications of this result are far-reaching, even if at this stage tentative. The solar dynamo appears to rely, for operation, on the subsurface flows to carry magnetic flux from the equatorial regions to the poles. Therefore a much better knowledge of solar structure near the surface which can be brought by more and more fully analysed and modelled helioseismic observations is likely to lead to better and more physical constraints on the causes of solar activity and its variations.

The data presented in Figs. 3–5 are intriguing, suggesting that important changes occurred in the Sun's behaviour early on during Cycle 23. However, important pieces of this puzzle are still missing. Among others, the alleged change in sunspot magnetic fields and temperatures claimed by Penn and Livingston (2006) should be revisited in the context of the statistical distribution of sunspot areas in various solar cycles; flare occurrence rates based on peak X-ray flux should be contrasted with changes in flare areas, H\_alpha flare index and CME rates, free magnetic energy, current helicity, etc. In view of such missing information it is probably premature to put forward a physical interpretation of these changes. Our objective here has been just to call attention to these puzzling findings and to place the papers in this volume in context.

# 5 Activity in Other Stars

While the study of solar activity is the primary topic of this issue of Space Science Reviews, it is worth recognising that the study of stellar activity is also a rich field of research. Of course, the observational means in stellar research are mostly fundamentally different and considerably more limited in capabilities from those used to study solar phenomena. Although severely limited by the observational difficulties, comparative studies of stellar and solar activity phenomena have been attempted to better understand the dynamo operating in the Sun and in the active stars.

Photometric observations of stars (changes in light output) have been used to deduce the presence of "starspots" on the stellar surface, implying strong magnetic fields and the operation of stellar dynamos (for a review, see Strassmeier 2009). Starspots are used to determine the rotation rate of stars, the photometric variability over intervals longer than the deduced rotation period is used to imply magnetic variability. This objective is generally difficult to achieve, as the continuous observation of target stars over their presumed activity cycles may require many years or even decades. Shorter-term photometric observations of starspots have been achieved. Such observations, together with modelling of the modulation of the photometric light curves has provided insights, if not certainties into the range of magnetic activity of different categories of solar-like stars. The potential for gaining enough knowledge into stellar activity is high and will increase as observational capabilities are developed, together with data analytical techniques which optimise the interpretation of what remains an incomparably smaller data set when compared to what is available for the Sun (Kővári and Oláh 2014, this issue).

Similarly to using helioseismology to study solar activity by studying changes to the internal structure of the Sun, asteroseismology or the detection of stellar oscillations has been used to study stellar activity. New results have been, to some extent, the by-product of the search for extra-solar planets. Indeed, the manifestations of stellar activity can be seen as a noise source for the detection of planets. The Kepler space mission which has observed a fixed field of a very large number of stars for a long interval first demonstrated that oscillations of solar-type stars can indeed be observed (Chaplin et al. 2010). A larger study identified up to 500 solar like stars with oscillations (Chaplin et al. 2011), a data set that can be used for the statistical study of stellar properties using the powerful methods that have proved so useful for the study of the solar interior. The question whether such asteroseismic observations can be interpreted in terms of stellar magnetic activity, however, is more complex and will require extensive modelling and the search for observations that can be used to constrain the models (Chaplin and Basu 2014, this issue).

Acknowledgements The Editors are greatly indebted to all the participants of the Workshop held in ISSI Bern on 11–15 November 2013 who brought their broad range of expertise and interest in solar physics to deepen our understanding of the issues related to the causes and consequences of the complex phenomenology of solar activity. We thank the staff of ISSI for their dedicated support: Prof. Rafael Rodrigo, Executive Director, and his colleagues Jennifer Fankhauser, Andrea Fischer, Saliba Saliba and Sylvia Wenger. The resulting collection of review papers was the outcome of the exchanges and fruitful collaboration among the participants; we thank them for their successful efforts to integrate the lessons learned in the different topics, as the reviews in the volume testify. Thanks are also due to the reviewers of the papers; in all cases the reviews were thorough and constructive and the volume bears witness to their contribution. Finally the Editors thank the staff of Space Science Reviews, as well as the production staff for their patience on occasion and for an excellently produced volume.

## References

- V. Archontis, Magnetic flux emergence and associated dynamic phenomena in the Sun. Philos. Trans. R. Soc. A 370, 3088–3113 (2012). doi:10.1098/rsta.2012.0001
- R. Arlt, N. Weiss, Solar activity in the past and the chaotic behaviour of the dynamo. Space Sci. Rev. (2014, this issue). doi:10.1007/s11214-014-0063-5
- M.J. Aschwanden, S.L. Freeland, Automated solar flare statistics in soft X-rays over 37 years of GOES observations: the invariance of self-organized criticality during three solar cycles. Astrophys. J. 754, 112 (2012). doi:10.1088/0004-637X/754/2/112
- H.D. Babcock, The Sun's polar magnetic field. Astrophys. J. 130, 364-365 (1959)

- W.T. Ball, Y.C. Unruh, N.A. Krivova, S. Solanki, T. Wenzler, D.J. Mortlock, A.H. Jaffe, Reconstruction of total solar irradiance 1974–2009. Astron. Astrophys. 541, A27 (2012). doi:10.1051/0004-6361/ 201118702
- A. Balogh, G. Erdős, The heliospheric magnetic field. Space Sci. Rev. 176, 177–215 (2013). doi:10.1007/ s11214-011-9835-3
- S. Basu, A.-M. Broomhall, W.J. Chaplin, Y. Elsworth, Thinning of the Sun's magnetic layer: the peculiar solar minimum could have been predicted. Astrophys. J. 758(1), 43 (2012), 6 pp. doi:10.1088/0004-637X/758/1/43
- G. Bazilevskaya, A.-M. Broomhall, Y. Elsworth, V.M. Nakariakov, A combined analysis of the observational aspects of the quasi-biennial oscillation in solar magnetic activity. Space Sci. Rev. (2014a, this issue). doi:10.1007/s11214-014-0068-0
- G.A. Bazilevskaya, E.W. Cliver, G.A. Kovaltsov, A.G. Ling, M.A. Shea, D.F. Smart, I.G. Usoskin, Solar cycle in the heliosphere and cosmic rays. Space Sci. Rev. (2014b, this issue). doi:10.1007/s11214-014-0084-0
- J.M. Borrero, K. Ichimoto, Magnetic structure of sunspots. Living Rev. Sol. Phys. 8, 4 (2011). http://www. livingreviews.org/lrsp-2011-4
- A.-M. Broomhall, W.J. Chaplin, Y. Elsworth, S.T. Fletcher, R. New, Is the current lack of solar activity only skin deep? Astrophys. J. 700, L162–L165 (2009). doi:10.1088/0004-637X/700/2/L162
- A.-M. Broomhall, P. Chatterjee, R. Howe, A.A. Norton, M.J. Thompson, The Sun's interior structure and dynamics, and the solar cycle. Space Sci. Rev. (2014, this issue). doi:10.1007/s11214-014-0101-3
- W.J. Chaplin, S. Basu, Inferences on stellar activity and stellar cycles from asteroseismology. Space Sci. Rev. (2014, this issue). doi:10.1007/s11214-014-0090-2
- W.J. Chaplin, Y. Elsworth, B.A. Miller, G.A. Verner, R. New, Solar p-mode frequencies over three solar cycles. Astrophys. J. 659, 1749–1760 (2007). doi:10.1086/512543
- W.J. Chaplin, T. Appourchaux, Y. Elsworth et al., The asteroseismic potential of Kepler: first results for solar type stars. Astrophys. J. Lett. 713, L169–L175 (2010). doi:10.1088/2041-8205/713/2/L169
- W.J. Chaplin, H. Kjeldsen, J. Christensen-Dalsgaard et al., Ensemble asteroseismology of solar-type stars with the NASA Kepler mission. Science 332, 213–216 (2011). doi:10.1126/science.1201827
- P. Charbonneau, Babcock–Leighton models of the solar cycle: questions and issues. Adv. Space Res. 39, 1661–1669 (2007)
- P. Charbonneau, Dynamo models of the solar cycle. Living Rev. Sol. Phys. 7, 3 (2010). doi:10.12942/lrsp-2010-3
- F. Clette, L. Svalgaard, J.M. Vaquero, E.W. Cliver, Revisiting the sunspot number: a 400-year perspective on the solar cycle. Space Sci. Rev. (2014, this issue). doi:10.1007/s11214-014-0074-2
- E.W. Cliver, The extended cycle of solar activity and the Sun's 22-yr magnetic cycle. Space Sci. Rev. (2014, this issue). doi:10.1007/s11214-014-0093-z
- I. Ermolli, K. Shibasaki, A. Tlatov, L. van Driel-Gesztelyi, Solar cycle indices from the photosphere to the corona: measurements and underlying physics. Space Sci. Rev. (2014, this issue). doi:10.1007/ s11214-014-0089-8
- L. Fletcher, B.R. Dennis, H.S. Hudson et al., An observational overview of solar flares. Space Sci. Rev. 159, 19–106 (2011). doi:10.1007/s11214-010-9701-8
- S. Forbush, Variation with a period of two solar cycles in the cosmic-ray diurnal anisotropy and the superposed variations correlated with magnetic activity. J. Geophys. Res. 74, 3451–3468 (1969)
- C. Frölich, Total solar irradiance observations. Surv. Geophys. 33, 453–473 (2012). doi:10.1007/s10712-011-9168-5
- G.E. Hale, On the probable existence of a magnetic field in sunspots. Astrophys. J. 28, 315–343 (1908)
- D.H. Hathaway, The solar cycle. Living Rev. Sol. Phys. 7, 1 (2010). doi:10.12942/lrsp-2010-1. http://www.livingreviews.org/lrsp-2010-1
- H. Hudson, L. Svalgaard, I. Hannah, Solar sector structure. Space Sci. Rev. (2014, this issue). doi:10.1007/ s11214-014-0121-z
- J.R. Jokipii, B. Thomas, Effects of drift on the transport of cosmic rays. IV. Modulation by a wavy interplanetary current sheet. Astrophys. J. 243, 1115–1122 (1981)
- J.R. Jokipii, E.H. Levy, W.B. Hubbard, Effects of particle drift on cosmic-ray transport. I. General properties, application to solar modulation. Astrophys. J. 213, 861–868 (1977)
- B.B. Karak, J. Jiang, M.S. Miesch, P. Charbonneau, A.R. Choudhuri, Flux transport dynamos: from kinematics to dynamics. Space Sci. Rev. (2014, this issue). doi:10.1007/s11214-014-0099-6
- R. Komm, The acoustic and magnetic solar cycle. Astron. Nachr. 331(9/10), 873–878 (2010). doi:10.1002/ asna.201011417
- J. Kota, J.R. Jokipii, Effects of drift on the transport of cosmic rays. VI. A three-dimensional model including diffusion. Astrophys. J. 265, 573–581 (1983)
- Z. Kővári, K. Oláh, Observing dynamos in cool stars. Space Sci. Rev. (2014, this issue). doi:10.1007/ s11214-014-0092-0

- R.B. Leighton, A magneto-kinematic model of the solar cycle. Astrophys. J. 156, 1–26 (1969)
- W. Livingston, M.J. Penn, L. Svalgaard, Decreasing sunspot magnetic fields explain unique 10.7 cm radio flux. Astrophys. J. Lett. 757, L8 (2012), 4 pp. doi:10.1088/2041-8205/757/1/L8
- I. Lopes, D. Passos, M. Nagy, K. Petrovay, Oscillator models of the solar cycle. Space Sci. Rev. (2014, this issue). doi:10.1007/s11214-014-0066-2
- K.G. McCracken, J. Beer, F.B. McDonald, Variations in the cosmic radiation, 1890–1986, and the solar and terrestrial implications. Adv. Space Res. 34, 397–406 (2004)
- A.A. Norton, P. Charbonneau, D. Passos, Hemispheric coupling: comparing dynamo simulations and observations. Space Sci. Rev. (2014, this issue). doi:10.1007/s11214-014-0100-4
- M.J. Penn, W. Livingston, Temporal changes in sunspot umbral magnetic fields and temperatures. Astrophys. J. Lett. 649, L45–L48 (2006). doi:10.1086/508345
- W.D. Pesnell, Solar cycle predictions. Sol. Phys. 281, 507–532 (2012). doi:10.1007/s11207-012-9997-5
- G.J.D. Petrie, K. Petrovay, K. Schatten, Solar polar fields and the 22-year activity cycle: observations and models. Space Sci. Rev. (2014, this issue). doi:10.1007/s11214-014-0064-4
- K. Petrovay, Solar cycle prediction. Living Rev. Sol. Phys. 7, 6 (2010). doi:10.12942/lrsp-2010-6. http:// www.livingreviews.org/lrsp-2010-6
- A.A. Pevtsov, Y.A. Nagovitsyn, A.G. Tlatov, A.L. Rybak, Long-term trends in sunspot magnetic fields. Astrophys. J. Lett. 742, L36 (2011), 4 pp. doi:10.1088/2041-8205/742/2/L36
- A.A. Pevtsov, M.A. Berger, A. Nindos, A.A. Norton, L. van Driel-Gesztelyi, Magnetic helicity, tilt, and twist. Space Sci. Rev. (2014, this issue). doi:10.1007/s11214-014-0082-2
- M.S. Potgieter, Solar modulation of cosmic rays. Living Rev. Sol. Phys. 10, 3 (2013). doi:10.12942/lrsp-2013-3. http://solarphysics.livingreviews.org/Articles/lrsp-2013-3/
- M. Rempel, Numerical sunspot models: robustness of photospheric velocity and magnetic field structure. Astrophys. J. 750, 62 (2012). doi:10.1088/0004-637X/750/1/62
- B. Schmieder, V. Archontis, E. Pariat, Magnetic flux emergence along the solar cycle. Space Sci. Rev. (2014, this issue). doi:10.1007/s11214-014-0088-9
- H. Schwabe, Sonnen-Beobachtungen im Jahre 1843. Astron. Nachr. 21(495), 233–236 (1844)
- E.J. Smith, The heliospheric current sheet. J. Geophys. Res. 106, 15819–15831 (2001)
- K.G. Strassmeier, Starspots. Astron. Astrophys. Rev. 17, 251–308 (2009)
- L. Svalgaard, The solar radio microwave flux (2009). http://wattsupwiththat.files.wordpress.com/2009/05/ svalgaard-solar-radio-flux.pdf
- L. Svalgaard, H.S. Hudson, The solar microwave flux and the sunspot number, in SOHO-23: Understanding a Peculiar Solar Minimum, ed. by S.R. Cranmer, J.T. Hoeksema, J.L. Kohl. ASP Conf. Ser., vol. 428 (Astronomical Society of the Pacific, San Francisco, 2010), p. 325
- K.F. Tapping, The 10.7 cm solar radio flux (F10.7). Space Weather **11**, 394–406 (2013). doi:10.1002/ swe.20064
- T. Thambyahpillai, H. Elliot, World-wide changes in the phase of the cosmic ray solar daily variation. Nature **171**, 918–920 (1953)
- M.J. Thompson, Helioseismology over the Solar Cycle, in SOHO-23: Understanding a Peculiar Solar Minimum, ed. by S.R. Cranmer, J. Todd Hoeksema, J.L. Kohl. ASP Conference Series, vol. 428 (Astronomical Society of the Pacific, San Francisco, 2010), pp. 23–28
- I.G. Usoskin, A history of solar activity over millennia. Living Rev. Sol. Phys. 10, 1 (2013). doi:10.12942/ lrsp-2013-1. http://www.livingreviews.org/lrsp-2013-1
- I.I. Virtanen, K. Mursula, North-south asymmetric solar cycle evolution: signatures in the photosphere and consequences in the corona. Astrophys. J. 781, 99 (2014), 8 pp. doi:10.1088/0004-637X/781/2/99
- Y.-M. Wang, Solar cycle variation of the Sun's low-order magnetic multipoles: heliospheric consequences. Space Sci. Rev. (2014, this issue). doi:10.1007/s11214-014-0051-9
- Y.-M. Wang, N.R. Sheeley Jr., Magnetic flux transport and the sun's dipole moment—new twists to the Babcock-Leighton model. Astrophys. J. 375, 761–770 (1991)
- W.R. Webber, M.S. Potgieter, R.A. Burger, A comparison of predictions of a wavy neutral sheet drift model with cosmic-ray data over a whole modulation cycle: 1976–1987. Astrophys. J. 349, 634–640 (1990)
- K.L. Yeo, N.A. Krivova, S.K. Solanki, Solar cycle variation in solar irradiance. Space Sci. Rev. (2014, this issue). doi:10.1007/s11214-014-0061-7