

SCIENTIFIC REPORTS

OPEN

Heartbeat of the Sun from Principal Component Analysis and prediction of solar activity on a millenium timescale

Received: 28 April 2015
Accepted: 25 September 2015
Published: 29 October 2015

V. V. Zharkova^{1,2,*}, S. J. Shepherd^{3,*}, E. Popova^{4,*} & S. I. Zharkov^{5,*}

We derive two principal components (PCs) of temporal magnetic field variations over the solar cycles 21–24 from full disk magnetograms covering about 39% of data variance, with $\sigma = 0.67$. These PCs are attributed to two main magnetic waves travelling from the opposite hemispheres with close frequencies and increasing phase shift. Using symbolic regression analysis we also derive mathematical formulae for these waves and calculate their summary curve which we show is linked to solar activity index. Extrapolation of the PCs backward for 800 years reveals the two 350-year grand cycles superimposed on 22 year-cycles with the features showing a remarkable resemblance to sunspot activity reported in the past including the Maunder and Dalton minimum. The summary curve calculated for the next millennium predicts further three grand cycles with the closest grand minimum occurring in the forthcoming cycles 26–27 with the two magnetic field waves separating into the opposite hemispheres leading to strongly reduced solar activity. These grand cycle variations are probed by $\alpha - \Omega$ dynamo model with meridional circulation. Dynamo waves are found generated with close frequencies whose interaction leads to beating effects responsible for the grand cycles (350–400 years) superimposed on a standard 22 year cycle. This approach opens a new era in investigation and confident prediction of solar activity on a millenium timescale.

Solar activity is manifested in sunspot occurrence on the solar surface characterized by the smoothed sunspot numbers, which were selected as a proxy of solar activity (see, for example, the top plot in <http://solarscience.msfc.nasa.gov/images/bfly.gif>). The sunspot numbers show quasi-regular maxima and minima of solar activity changing approximately every 11 years, with changing leading magnetic polarity in a given hemisphere (or 22 years for sunspots with the same polarity) reflecting changing magnetic activity of the Sun¹.

The longest direct observation of solar activity is the 400-year sunspot-number series, which depicts a dramatic contrast between the almost spotless Maunder and Dalton minima, and the period of very high activity in the most recent 5 cycles^{2,3}, prior to cycle 24. Many observations indicate essential differences between the activity occurring in the opposite hemispheres for sunspots⁴ and for solar and heliospheric magnetic fields⁵.

Prediction of a solar cycle through sunspot numbers has been used for decades as a way of testing accuracy of solar dynamo models, including processes providing production, transport and disintegration of the solar magnetic field. Cycles of magnetic activity are associated with the action of a dipole

¹Northumbria University, Department of Mathematics & Information Sciences, Newcastle upon Tyne, NE2 1XE, UK.

²Institution of Space Science Research, Space Physics Department, Kiev, 03022, Ukraine. ³University of Bradford, School of Engineering, Bradford, BD7 1DP, UK. ⁴Skobeltsyn Institute of Nuclear Physics, Moscow 119234, Russia.

⁵University of Hull, Department of Physics and Mathematics, Kingston upon Hull, HU6 7RX, UK. *These authors contributed equally to this work. Correspondence and requests for materials should be addressed to V.V.Z. (email: valentina.zharkova@northumbria.ac.uk)

solar dynamo mechanism called ‘ $\alpha-\Omega$ dynamo’⁶. It assumes the action of solar dynamo to occur in a single spherical shell, where twisting of the magnetic field lines (α -effect) and the magnetic field line stretching and wrapping around different parts of the Sun, owing to its differential rotation (Ω -effect), are acting together^{7,8}.

As a result, magnetic flux tubes (toroidal magnetic field) seen as sunspots are produced from the solar background magnetic field (SBMF) (poloidal magnetic field) by a joint action of differential rotation (Ω -effect) and radial shear (α -effect), while the conversion of toroidal magnetic field into poloidal field is governed by the convection in the rotating body of the Sun. The action of the Coriolis force on the expanding, rising (compressed, sinking) vortices results in a predominance of right-handed vortices in the Northern hemisphere and left-handed vortices in the Southern hemisphere leading to the equatorward migration of sunspots during a solar cycle duration visible as butterfly diagrams (see <http://solarscience.msfc.nasa.gov/images/bfly.gif>, the bottom plot).

The last few decades were extremely fruitful in investigating the contribution of various mechanisms to the dynamo processes including the conditions for dynamo wave generation from the mean dynamo models with different properties of solar and stellar plasmas, as discussed in the recent reviews^{7,8}.

As usual, the understanding of solar activity is tested by the accuracy of its prediction. The records show that solar activity in the current cycle 24 is much lower than in the previous three cycles 21–23 revealing more than a two-year minimum period between cycles 23 and 24. This reduced activity in cycle 24 was very surprising because the previous five cycles were extremely active and sunspot productive forming the Modern Maximum^{2,3}. Although the reduction of solar activity in cycle 24 led some authors to suggest that the Sun is on its way towards the Maunder Minimum of activity⁹.

However, most predictions of solar activity by various methods, such as considering linear regression analysis¹⁰, neural network forecast¹¹, or a modified flux-transport dynamo model calibrated with historical sunspot data from the middle-to-equator latitudes¹², anticipated a much stronger cycle 24¹⁰. There were only a few predictions of the weaker cycle 24¹³ obtained with the high diffusivity Babcock-Leighton dynamo model applied to polar magnetic fields as a new proxy of solar activity. However, a dynamo model with a single wave was shown to be unable to produce reliable prediction of solar activity for longer than one solar cycle because of the short memory of the mean dynamo¹⁴.

Consistent disagreement between the sunspot numbers, measured averaged sunspot numbers and the predicted ones by a large number of complex mathematical models for cycle 24, is undoubtedly the result, which emphasizes the importance of different physical processes occurring in solar dynamo and affecting complex observational appearance of sunspots on the surface.

Results

Two principal components as two dynamo waves. In order to reduce dimensionality of these processes in observational data, Principal Component Analysis (PCA) was applied¹⁵ to low-resolution full disk magnetograms captured by the Wilcox Solar Observatory¹⁶. This approach revealed a set of more than 8 independent components (ICs), which seem to appear in pairs¹⁵, with two principal components (PCs) covering about 39% of the variance of the whole magnetic field data, or standard deviation of $\sigma = 0.67$. The main pair of PCs is associated with two magnetic waves of opposite polarities attributed to the poloidal field produced by solar dynamo from a dipole source¹⁷.

The two principal components (PCs) derived from solar background magnetic field (SBMF)¹⁵ (cycle 21–23) and predicted for cycle 24–26 are presented in Fig. 1 (the upper plot). For the first time PCA allowed us to detect, *two magnetic waves* in the SBF¹⁵ and not a single one assumed in the mean dynamo models. These waves are found originating in the opposite hemispheres and travelling with an increasing phase shift to the Northern hemisphere in odd cycles and the Southern hemisphere in even cycles¹⁵. This can explain the well-observed North-South asymmetry in sunspot numbers, background magnetic field, flare occurrences and so on (see Zharkov *et al.*⁴ and references therein) defining the active hemisphere for odd (North) and even (South) cycles.

The formation of magnetic flux tubes emerging on the solar surface as sunspots can be considered as a result of interaction in the solar interior of the two magnetic waves of the solar background magnetic field¹⁵ when their phase shift is not very large. These two magnetic waves of the poloidal field can account for the observed sunspot magnetic field¹⁸, or averaged sunspot numbers, after their amplitudes are added together into the summary wave (Fig. 1, bottom plot) and converted to the modulus curve by taking modulus of the summary curve¹⁹ (Fig. 2, bottom plot). The modulus curve plotted for cycles 21–23 in Fig. 2 (top plot) corresponds rather closely to the averaged sunspot numbers for cycles 21 and 22 while being noticeably lower than the sunspot curve for cycle 23, which anticipated the recently discovered sunspot calibration errors occurred in the past few decades²⁰.

The maximum (or double maximum for the waves with a larger phase shift of solar activity for a given cycle) coincides with the time when each of the waves approaches a maximum amplitude and the hemisphere where it happens becomes the most active one. This can account naturally for the north-south asymmetry of solar activity often reported in many cycles. Also the existence of two waves in the poloidal magnetic field instead of a single one, used in most prediction models, and the presence of a variable phase difference between the waves can naturally explain the difficulties in predicting sunspot activity on a scale longer than one solar cycle with a single dynamo wave¹⁴ since the sunspot activity is associated with the modulus summary curve of the two dynamo waves¹⁹ that is a derivative from these two waves.

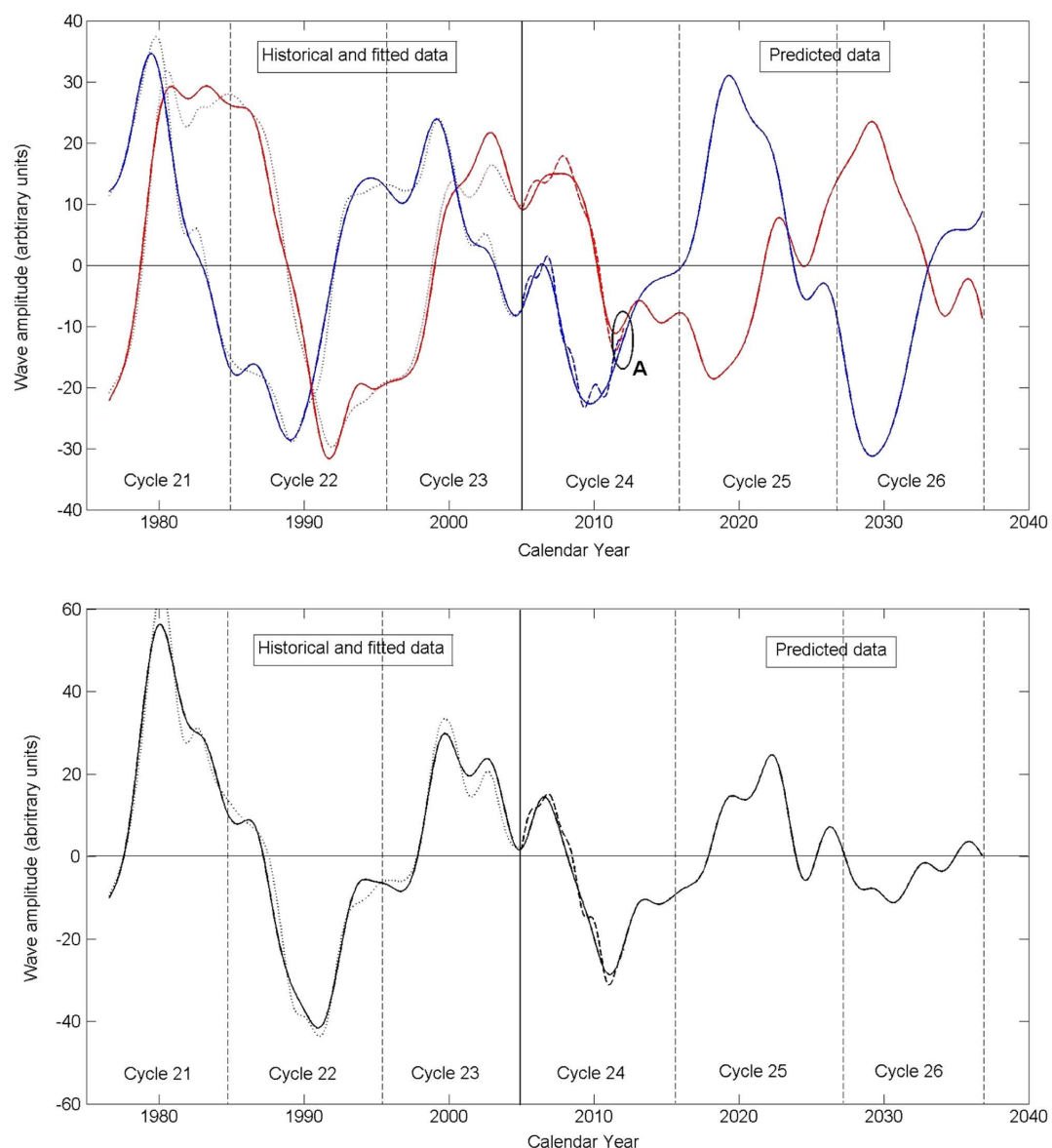


Figure 1. Top plot: the two principal components (PCs) of SBF (blue and red curves) obtained for cycles 21–23 (historic data¹⁵) and predicted¹⁹ for cycles 24–26 with the Eqs. (2)–(3). The dotted lines show the PCs derived from the data and the solid lines present the curves plotted from formulae 2 (blue) and 3 (red). The accuracy of fit of the both PC curves is better than 97%. The point A shows the current time. The cycle lengths (about 11 years) are marked at the minima by the vertical lines. The bottom plot: The summary PC derived from the two PCs above for the ‘historical’ (21–23) and predicted cycles (24–26) data. The dotted curve shows PCs derived from the data and the solid line - from the solid curves from the top plot using formulae 2–3. The cycle lengths (about 11 years) are again marked by the vertical lines at the cycle minima. All the plots are a courtesy of Shepherd *et al.*¹⁹. © AAS. Reproduced with permission.

Mathematical description of the observed magnetic waves. Amplitude and frequency variations of these waves, or PCs, over time are found using symbolic regression analysis²¹ with Euriqa software (see the Methods section for data analysis¹⁹). The wave amplitudes follow the product of two cosine functions ($\cos * \cos$), while the frequencies follow a nested function ($\cos(\cos)$) depicting the fact that the waves periodically change their frequency and phase with time. These formulae are used to extract the key parameters of the principal components of SBF waves, which are, in turn, used for prediction of the overall level of solar activity for solar cycles 24–26 associated with the averaged sunspot numbers¹⁹. The accuracy of these formulae for prediction of the principal components is tested for cycle 24 showing the predicted curve fitting very closely (with an accuracy of about 97.5%) the PCs derived from the observations of SBF and sunspot numbers¹⁹.

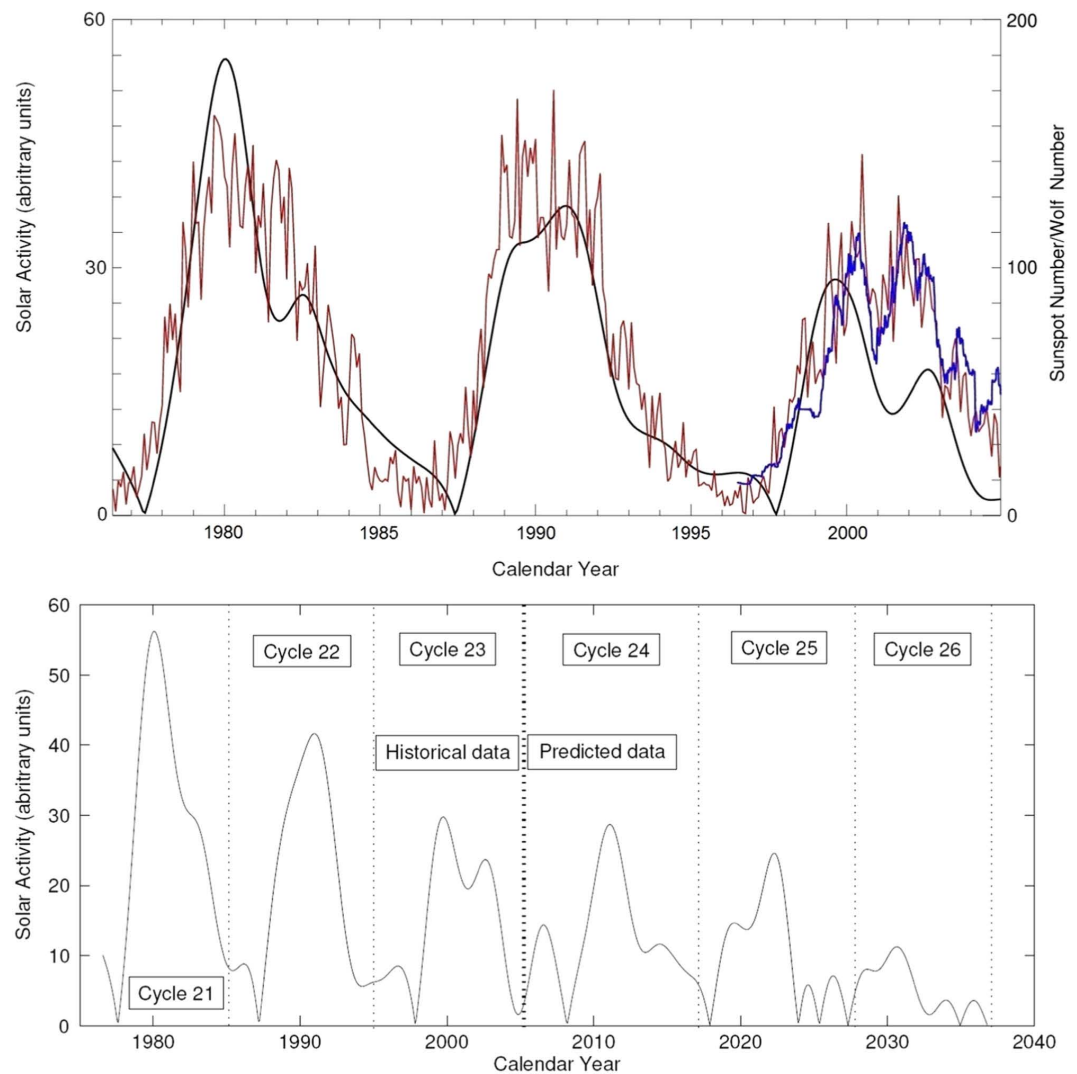


Figure 2. Top plot: Comparison of the modulus summary curve (black curve) obtained from the summary curve in Fig. 1 with averaged sunspot numbers (brown curve) and magnetic field (blue curve) for cycles 21–23. Bottom plot: The modulus summary curve associated with the sunspot numbers derived for cycles 21–23 (plotted in the top plot) and calculated for cycles 24–26 using the mathematical formulae (2–3). The plots are a courtesy of Shepherd *et al.*¹⁹. © AAS. Reproduced with permission.

For the forthcoming cycles 25 and 26 (Fig. 1) the two waves are found to travel between the hemispheres with decreasing amplitudes and increasing phase shift approaching nearly a half period in cycle 26. This leads, in fact, to a full separation of these waves in cycle 26 into the opposite hemispheres¹⁹. This separation reduces any possibility for wave interaction for this cycle that will result in significantly reduced amplitudes of the summary curve and, thus, in the strongly reduced solar activity in cycle 26¹⁹, or the next Maunder Minimum⁹ lasting in 3 cycles 25–27.

Prediction of solar activity on millennium scale. By far the most impressive achievement to-date of this approach is its ability to make very long term predictions of solar activity with high accuracy over the timescales of many centuries. The summary curve of the two principal components (magnetic waves) expressed by the formulae (2 and 3) in the Method of data analysis¹⁹ is calculated backwards and forwards for the period 1200–3200 years as shown in Fig. 3.

Remarkably, our current prediction of the summary curve backwards by 800 years shown in the left (from oval) part of Fig. 3, corresponds very closely to the sunspot data observed in the past 400 years as indicated by the brackets in Fig. 3, with the black oval marking the data used to derive Eq. (2) and (3) defining the wave variations. We predict correctly many features from the past, such as: 1) an increase in solar activity during the Medieval Warm period; 2) a clear decrease in the activity during the Little Ice

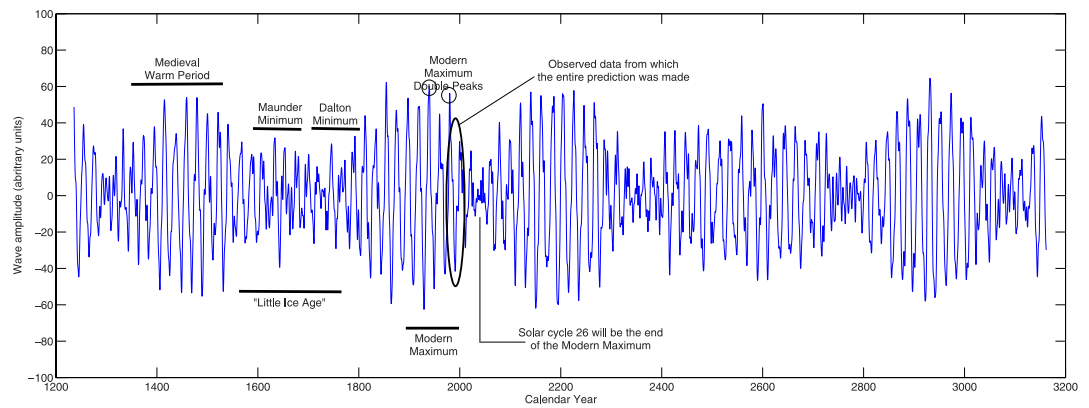


Figure 3. The predicted summary wave (the sum of two principal components) calculated from 1200 to 3200 years from the ‘historical’ period (cycles 21–23) marked with a black oval. The historical maxima and minima of the solar activity in the past are marked by the horizontal brackets.

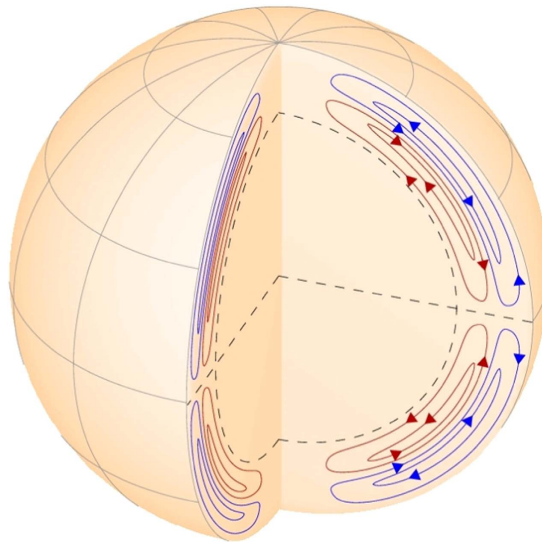


Figure 4. The schematic dynamo model with two cells in the solar interior having the opposite meridional circulation as derived from HMI/SDO observations by Zhao *et al.*²⁴. © AAS. Reproduced with permission.

Age, the Maunder Minimum and the Dalton Minimum; 3) an increase in solar activity during a modern maximum in 20th century.

This visual correspondence in the features between the summary curve and the averaged sunspot numbers is most surprising, given the fact that the principal components are derived from the solar background magnetic field, and they are not linked directly to the sunspot numbers (see Methods for data analysis) besides the modulus summary curve derived from the principal components as shown in Fig. 2.

The summary curve reveals a superposition of the amplitudes of the two dynamo waves, or a ‘beating’ effect creating two resulting waves: one of higher frequency (corresponding to a classic 22-year cycle) and a second wave of lower frequency (corresponding to a period of about 350–400 years), which modulates the amplitude of the first wave. It appears that this grand cycle has a variable length from 320 years (in 18–20 centuries) to 400 (in 2300–2700) predicted for the next millennium. Amplitudes in the shorter grand cycles are much higher than the amplitudes in the longer ones.

This long-term ‘grand’ cycle was previously postulated in 1876 by Clough²² as a 300-year cycle superimposed on the 22 year cycle using the observations of aurorae, periods of grape harvests etc, which was later suggested to have a period of about 205 years²³. These periods are close to those reported for the last 800 years in the summary curve plotted in Fig. 3 derived from the observed magnetic field variations.

The spectacular accuracy of the historical fit in the past 800 years gave us the confidence to extrapolate the data into the future for a similar epoch of 1200 years (Fig. 3, right part of the curve) clearly

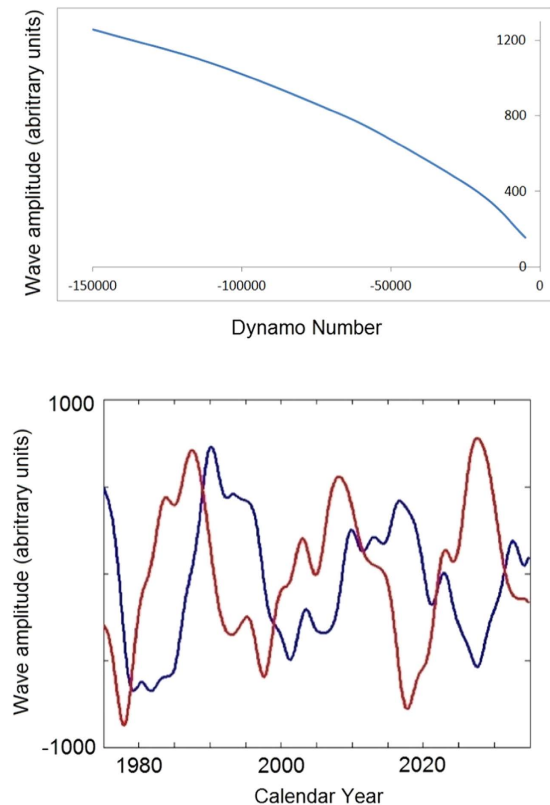


Figure 5. Top plot: Dependence of the solar dynamo-number $D = R_\alpha R_\Omega$ on a magnitude of the toroidal magnetic field (for details of the parameters see the text). Bottom plot: Variations of the toroidal magnetic field simulated for cycles 21–26 with two layer $\alpha\Omega$ dynamo model (see Methods section) for the inner (red line) and upper (blue line) layers. One arbitrary unit corresponds to 1–1.5 Gauss (see text for details).

showing, as expected, several 350–400-year grand cycles. We note, in particular, a decreasing activity for solar cycles 25 and 26 coinciding with the end of the previous 350–400-year grand cycle and then increase of the solar activity again from cycle 27 onwards as the start of a new grand cycle with an unusually weak cycle 30. Hence, cycles 25–27 marks a clear end of the modern grand period that can have significant implications for many aspects of solar activity in human lives including the current debate on climate change.

Discussion

Preliminary interpretation with the two layer $\alpha - \Omega$ dynamo model. Now let us attempt some preliminary interpretation of the two principal components, or two magnetic waves of solar poloidal field, generated by the solar dynamo in two different cells, similar to those derived by Zhao *et al.*²⁴ from helioseismological observations (Fig. 4), in order to fit the background magnetic field observations (Figs 1 and 3). This can be achieved with the modified Parker’s non-linear two layers dynamo model for two dipoles¹⁷ with meridional circulation: in the layer 1 of the top cell and layer 2 of the bottom cell from Fig. 4 (see Methods section for the model description) tested for the interpretation of latitudinal waves in the solar background magnetic field for cycles 21–23¹⁷ derived with PCA¹⁵.

The simulation results presenting the toroidal magnetic field are plotted in Fig. 5 (bottom plot) derived from the poloidal field (Fig. 1, top plot) for a period of six 11-year cycles using the dynamo equations (16–19) from Popova *et al.*¹⁷. The curves for poloidal (derived with PCA) and toroidal fields (simulated with the dynamo model) are found to have similar periods of oscillations whilst having opposite polarities (or having the phase shift of a half of the period), being in anti-phase every 11 years as previously reported^{4,25}. The amplitude of generated toroidal magnetic field is plotted versus the dynamo number $D = R_\alpha \cdot R_\Omega$ in Fig. 5 (top plot).

Furthermore, in cycles 25–27 and, especially, in cycle 26, the toroidal magnetic field waves generated in these two layers become fully separated into the opposite hemispheres, similar to the two PC waves attributed to poloidal field (Fig. 1, top plot), that makes their interaction minimal. This will significantly reduce the occurrence of sunspots in any hemisphere, that will result in a very small solar activity index for this cycle, resembling the Maunder Minimum occurred in the 17th century.

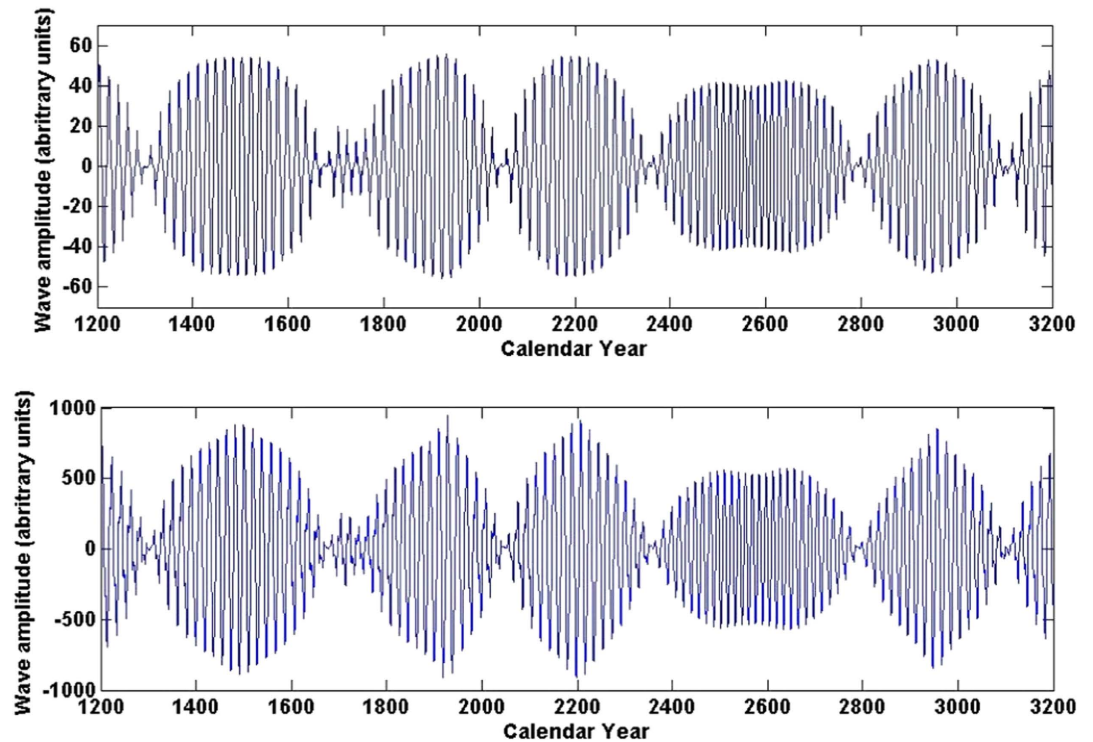


Figure 6. Variations of the summary poloidal (top plot) and toroidal (bottom plot) magnetic fields simulated for 2000 years with the two layer $\alpha\Omega$ -dynamo model (see Methods section) with the parameters derived from the two PCs from Fig. 1 using mathematical formulae (2–3). One arbitrary unit corresponds to 1–1.5 Gauss (see text for details).

Using the same dynamo parameters derived from the observed principal components for these 6 cycles, let us extend the calculation (see the Methods for details) to a longer period of two millennia shown in Fig. 6 for both poloidal (top plot) and toroidal (bottom plot) fields. According to the dynamo theory and analysis of observational data^{7,27} the generated toroidal field is much stronger than the poloidal. Although, exact values of the amplitudes of these fields in the solar convection zone are unknown and estimated from dynamo models. In our simple model the amplitude of toroidal field at the maximum is about 1000 Gauss, and of the poloidal one is of the order of several tens of Gauss. Hence, in Figs 5 and 6 one arbitrary unit approximately corresponds to 1–1.5 Gauss.

It can be seen that variations of the model magnetic fields (Fig. 6) generated by the two dipole sources located in different layers reproduce the main features discovered in Fig. 3, e.g. modulation of the amplitude of 22 year cycle by much slower oscillations of about 350 years, different duration (320–400) and amplitudes of different grand cycles. These variations are governed by different dynamo parameters as discussed below.

Beating effect of two dynamo waves with close frequencies. The waves generated by a dynamo mechanism in each layer are found to have similar (but not equal) frequencies caused by a difference in the meridional flow amplitudes in the two layers (Fig. 5, bottom plot). In order to reproduce the summary curve in Fig. 3 from the two original waves, or PCs, the dynamo waves generated in different layers with an amplitude A_0 have to have close but not equal frequencies ω_1 and ω_2 (or periods varying between 20 and 24 years), similar to Gleissberg's cycle^{7,26}.

The interference of these waves enabled by diffusion of the waves in the solar interior from the bottom to the top layer²⁷ leads to formation of the resulting envelope of waves $Y(t)$, or beating effect (see Fig. 3 and theoretical plots in Fig. 6), showing oscillations of a higher frequency $(\omega_1 + \omega_2)/2$ within the envelope and those of the envelope itself with a lower frequency of $(\omega_1 - \omega_2)/2$ (or in a *grand cycle*) as follows:

$$Y(t) \approx \left| \frac{2A_0}{k(\omega_1^2 - \omega_2^2)} \sin\left(\frac{\omega_1 - \omega_2}{2}t\right) \right| \sin\left(\frac{\omega_1 + \omega_2}{2}t\right), \quad (1)$$

where k is some parameter defining properties of the solar interior where the waves propagate, e.g. diffusivity, dynamo number (α and Ω effects) and meridional circulation.

Frequency and period variations. The beating effect between these frequencies can easily explain seemingly sporadic variations of high frequency amplitudes and the period of the low-frequency envelope wave in the resulting grand cycles seen in both the observational curve (Fig. 3 and theoretical curves (Fig. 6) reproducing the observational one. The higher the difference of frequencies the larger is the frequency, or a shorter period, of the grand cycle (350 years) and the smaller is a number of high frequency waves (≈ 22 year period) within this grand cycle. This effect is clearly seen in Figs 3 and 6, where the grand periods with a lower number of 22 year cycles are shorter (300–340 years, 2nd, 3rd and 5th grand cycles in Fig. 3), while those with higher number of 22-year cycles are longer (360–400 years, the 1st and 4th in Fig. 3).

The difference in frequencies of the dynamo waves in two layers is governed by the variations of velocities of meridional circulations in the very top and the very bottom zones of these two layers (see the Method section) (schematically presented in Fig. 4 from Zhao *et al.*²⁴). The frequency of a wave is reduced (or its period is increased) when the meridional circulation has higher velocities and this frequency is increased (or its period is decreased) when the meridional circulation is slower. It means that the meridional circulation acts as a drag force for dynamo waves generated in each layer altering their natural frequencies that would occur without the circulation.

For example, within the two layers model considered, and taking into account that the low frequency cycles can have length T_g from 20 to 24 years (variations within Gleissberg's cycle⁷), in order to produce the grand cycle with a beating period of 350 years, the periods of the dynamo waves in two layers should vary as follows: for the sunspot activity period $T_g = 20$ years - for the inner layer wave $1 - T_1 = 18.9$ years (corresponding to the velocity of meridional circulation about $V = 7-8$ m/s), for the upper layer wave $2 - T_2 = 21$ years ($V = 9-10$ m/s); for the activity period $T_g = 24$ years: the inner layer wave $1 - T_1 = 22.46$ years ($V = 10-11$ m/s), the upper layer wave $2 - T_2 = 25.8$ years ($V = 13-14$ m/s).

If the grand cycle is 400 years, then the dynamo wave periods in two layers would slightly change; e.g. for the cycle period $T_g = 20$ years - for the inner layer wave $1 - T_1 = 19$ years ($V = 7-8$ m/s), for the upper layer wave $2 - T_2 = 21$ years ($V = 9-10$ m/s); for the period of $T_g = 24$ years: the inner layer wave $1 - T_1 = 22.6$ years ($V = 10-11$ m/s), the upper layer wave $2 - T_2 = 25.53$ years ($V = 13-14$ m/s).

It can be seen that the period of the wave 1 generated in the inner layer (at the bottom of the convective zone) remains more or less stable at about $T_1 = 19$ years (for generation of the low frequency activity period $T_g = 20$ years) or $T_1 = 22.6$ year (for $T_g = 24$ years). While the period of the wave 2 generated in the upper layer should have larger fluctuations (e.g. $T_2 = 25.8$ years for 350 grand cycle versus $T_2 = 25.53$ years for 400 years grand cycle). These fluctuation are likely to be affected by the physical conditions in the solar interior, where the wave 2 is formed and the wave 1 has to travel through and to interact with the wave 2 to cause the beating effect combining the grand (ranging in 300–400 years) and short (ranging in 20–24 years) cycles seen in Fig. 3 as reproduced with the dynamo model in Fig. 6 for both poloidal and toroidal magnetic fields.

Of course, estimations of the wave beating above are rather preliminary, given the fact that the PCs (or dynamo waves) in each layers comprise at least 5 waves with close frequencies as discussed in the Method section (Eqs. 2 and 3). This results in much more complex beating effects derived from PCA as presented in Fig. 3. The dynamo calculations only partially reproduced the long cycle with a period of about 350 years, which is the same for the whole millennium. However, in order to reproduce the full summary curve with the variable long-term period in Fig. 3 more detailed dynamo simulations including quadruple magnetic sources in all the three layers (shown in Fig. 4) are required.

Wave amplitude variations. The amplitudes of dynamo waves are affected by the variations of both α and Ω effects, or by the dynamo number D , i.e. a decrease of the negative dynamo number D (or its increase in absolute value) leads to an increase of toroidal field amplitude (see Fig. 5, top plot).

This effect can be observed in both the observational (Fig. 3) and theoretical (Fig. 6) plots. In shorter grand cycles (with periods of 300–340 years), e.g. in 1800–2000 years and 2100–2350 years, the amplitudes of the high frequency wave ($T_g = 20-24$ years) are much higher than in longer cycles (periods of 350–400 years) in 1300–1650 years or 2400–2800 years. Although, in order to reproduce more closely the whole variety of observational features on a longer timescale, more detailed 3D model simulations are required.

Therefore, the derived mathematical laws in cyclic variations of principal components of the observed solar magnetic field, which fit closely most of the observational features of solar activity in the past as shown in Fig. 3 and reproduced by the dynamo model in Fig. 6 opens a new era in the investigation of solar activity on millennium scale. By combining the observational curve with simulations of solar dynamo waves in two layers, it is possible to derive better understanding of the processes governing solar activity and produce long-term prediction of solar activity with impressive accuracy.

Methods

Derivation of parameters of the observed magnetic waves. In order to distill the main parameters of the waves present in the observational solar magnetic data, one needs to reduce their

dimensionality with the Principal Component Analysis (PCA)²⁸. PCA is an orthogonal linear transformation allowing a vector space to be transformed to a new coordinate system, reducing the multi-dimensional data to lower dimensions for analysis, so that the greatest variance by any projection of the data lies on the first coordinate called the Principal Component (PC) with the second PC orthogonal to the first is defined by the second largest variance. This technique simultaneously (i) reduces the data dimensionality, (ii) increases the signal-to-noise ratios and (iii) orthogonalises the resulting components so that they can be ascribed to separate physical processes (see Zharkova *et al.*¹⁵ for more details). The PCA is an exact method, and its accuracy defined only by the noise of measurements, σ_x^2 , of the original vector²⁹.

PCA was applied to low-resolution full disk solar background magnetic field (associated with the poloidal magnetic field) only become available from cycle 21 to cycle 24 as measured by the Wilcox Solar Observatory (with accuracy better than 0.5 Gauss, or the measurement error $\sigma_x^2 = 0.0025$). We derive the dominant eigenvalues (0.1 and 1.0) covering the maximum variance of 39%¹⁵ defining the eigenfunctions, or Principal Components (PCs), which came as a pair of waves. These PCs are considered as the main (dipole) dynamo waves of the solar poloidal magnetic field.

By applying a 3-year running averaging filter, any short-term (<3 years) fluctuations of magnetic field data are removed allowing us to keep the accuracy of PCA not worse than the measurement error (Wentzell and Lohnes³⁰). The overall PCA accuracy of defining its eigen values from the WSO data with known measurement error (see Faber *et al.*²⁹) is not worse than 0.2%. Running PCA on a combination of magnetic field measurements for any two cycles, or for all four cycles 21–24 produces, within the error of 0.2%, the same eigenvalues as for the three cycles used in PCA¹⁵.

For classification of the derived PCs we apply the symbolic regression approach based on the Hamiltonian principle implemented in the *Eurika* software²¹. This allows us to derive the exact mathematical formulae for the amplitude variations and phase shifts of both principal components as follows¹⁹:

for wave 1:

$$f(t) = \sum_{k=1}^N A_k \cos(\omega_{k,1}t + \phi_{k,1}) \cos(B_k \cos(\omega_{k,2}t + \phi_{k,2})), \quad (2)$$

for wave 2:

$$f(t) = \sum_{k=1}^N C_k \cos(\omega_{k,3}t + \phi_{k,3}) \cos(D_k \cos(\omega_{k,4}t + \phi_{k,4})), \quad (3)$$

where the parameters with ω define the corresponding wave frequencies and ϕ define their phase shifts. Shepherd *et al.*¹⁹ found that the approximations with only $N=5$ terms in the series above allow them to capture the functions describing the waves of PCs for the cycles 21–24 with an accuracy better than 97%¹⁹. As expected, any attempts to distill the parameters from the original magnetic field data (before deriving PCs) were unsuccessful indicating the very complex nature of the original magnetic field waves.

These two PCs are used for calculation of the summary wave (a sum of amplitudes) and the modulus summary wave (reflected to the positive amplitudes only) linked to the averaged sunspot numbers currently used for definition of solar activity.

Non-linear $\alpha\Omega$ dynamo model in a two-layer medium with meridional circulation. In order to understand the basic features of the derived PCs, let us use Parker's $\alpha\Omega$ -dynamo model with two layers with meridional circulation¹⁷ updated by considering a non-linear dynamo process. It is assumed that dynamo waves are generated by the dipole sources only located in two layers: one dipole in the sub-surface layer and the other dipole deeply in the solar convection zone (see Fig. 4); and the parameters (dynamo number and meridional circulation) of magnetic field generation in each layer are different¹⁷.

This results in the simultaneous existence of two magnetic waves with different periods and phase shifts¹⁷, similar to those derived with PCA (see Fig. 1). For the sake of simplicity this approach excludes the dynamo waves generated by quadruple sources in both layers accounting for the other six independent components¹⁷, which are shown to slightly modify the overall appearance of magnetic waves that will be considered in the forthcoming paper.

The dynamo equations describing the generation and evolution of the solar magnetic field in a two-layer medium, are obtained from a system of electrodynamic equations for mean fields in the assumption that the dynamo wave propagates in a thin spherical shell¹⁷. In this case, the magnetic field is averaged along the radius within a certain spherical shell and the terms describing the curvature effects near the pole are excluded. In addition, in this approximation, we assume that the magnetic field is generated independently in either layer. As a result, the equations take the form of equations (16–19) by Popova *et al.*¹⁷ solved numerically using the method of lines³¹ and verified analytically by the low-mode approach³².

In these equations the dynamo number $D = R_\alpha R_\Omega$ is defined by the parameters R_α and R_Ω describing, respectively, the intensity of α -effect, and the differential rotation, or Ω -effect. The latitudinal profile of the poloidal magnetic field is assumed for simplicity to be proportional $\cos(\theta)$ where θ is the solar latitude measured from the equator.

We consider the α -effect the amplitude $F(t, \text{coordinates})$ and widely used algebraic quenching⁷ in a form:

$$\alpha = \frac{F(t, \text{coordinates})}{1 + \xi^2 (B_{\text{toroidal}})^2} = R_\alpha \frac{\alpha_0(\theta)}{1 + \xi^2 (B_{\text{toroidal}})^2}. \quad (4)$$

For calculations the amplitude of α effect, R_α is moved to the dynamo number D , while the algebraic quenching of the helicity is used for stabilization of a magnetic field growth, e.g. redefining $\alpha = \frac{\alpha_0(\theta)}{(1 + \xi^2 B^2)}$, where $\alpha_0(\theta) = \cos \theta$ is the helicity in unmagnetized medium and $B_0 = \xi^{-1}$ is the magnetic field, for which the α -effect is considerably suppressed.

The contribution of differential rotation into the generation of magnetic field is defined²⁷ by the terms $\cos(\theta) \frac{\partial A}{\partial \theta}$ for one layer or $\cos(\theta) \frac{\partial a}{\partial \theta}$ for another layer, following the general trend of being maximum at the equator and minimal at the poles. The dynamo number $D = R_\alpha R_\Omega$ also includes an amplitude of differential rotation R_Ω , which can vary in different layers.

Since the frequencies of magnetic waves generated by dynamo mechanism are known to be mostly affected by meridional circulation velocities^{17,32} while their amplitudes are governed by the variations of dynamo number D , then the PC waves can be reproduced in different layers with different dynamo numbers and slightly different meridional circulation velocities.

In each layer we consider a 1D dynamo model with meridional circulation V being a function dependent only on θ e.g. $V = v \sin(2\theta)$, so that it vanishes at the poles and is maximal at the middle latitudes approaching amplitudes of 9–15 m/s^{24,33}. Also, to comply with the material conservation rule, the meridional circulation the multi-cellular meridional circulation has to have the opposite directions in upper and inner layers of the cells in as shown in Fig. 4 suggested earlier by Dikpati^{34,35} and Popova *et al.*¹⁷ that was recently confirmed from the helioseismic observations with HMI by Zhao *et al.*²⁴.

In general, there are three layers (see Fig. 4), in which the meridional circulation affects the magnetic field: 1- the very top layer in the upper cell and 2- the very bottom layer in the inner cell where the meridional circulation has the same direction but different velocities and 3- the middle layer where cells have a boundary and their the circulation has the opposite direction, complying with the mass conservation law. In the current model we consider the top and bottom layers (1–2) only to reproduce the long-term oscillations produced in them, while oscillations in the middle layer 3 affecting the short-term biennial³² are out of scope of the current study.

For each layer the principal components (PCs) of poloidal magnetic field were substituted into the dynamo equations (17 and 19) of Popova *et al.*¹⁷, from which the corresponding toroidal magnetic field components are derived fitting these PCs. Then we substitute the toroidal and poloidal components into the dynamo equations (16 and 18) of Popova *et al.*¹⁷ for corresponding layers and derive their dynamo numbers D .

References

- Hathaway, D. H., Wilson, R. M. & Reichmann, E. J. Group Sunspot Numbers: Sunspot Cycle Characteristics. *Sol. Phys.* **211**, 357–370 (2002).
- Solanki, S. K., Usoskin, I. G., Kromer, B., Schüssler, M. & Beer, J. Unusual activity of the Sun during recent decades compared to the previous 11,000 years. *Nature* **431**, 1084–1087 (2004).
- Solanki, S. K. & Krivova, N. A. Analyzing Solar Cycles. *Science* **334**, 916 (2011).
- Zharkov, S., Gavryuseva, E. & Zharkova, V. The Observed Long- and Short-Term Phase Relation between the Toroidal and Poloidal Magnetic Fields in Cycle 23. *Sol. Phys.* **248**, 339–358 (2008).
- Bravo, S. & González-Esparza, J. A. The north-south asymmetry of the solar and heliospheric magnetic field during activity minima. *Geophys. Res. Lett.* **27**, 847–849 (2000).
- Parker, E. N. Hydromagnetic Dynamo Models. *Astrophys. J.* **122**, 293 (1955).
- Brandenburg, A. & Subramanian, K. Astrophysical magnetic fields and nonlinear dynamo theory. *Phys. Rep.* **417**, 1–209 (2005).
- Jones, C. A., Thompson, M. J. & Tobias, S. M. The Solar Dynamo. *Space Sci. Rev.* **152**, 591–616 (2010).
- Lockwood, M., Owens, M. J., Barnard, L., Davis, C. J. & Steinhilber, F. The persistence of solar activity indicators and the descent of the Sun into Maunder Minimum conditions. *Geophys. Res. Lett.* **38**, 22105 (2011).
- Pesnell, W. D. Predictions of Solar Cycle 24. *Sol. Phys.* **252**, 209–220 (2008).
- Maris, G. & Onćica, A. Solar Cycle 24 Forecasts. *Sun and Geosphere* **1**, 010000–11 (2006).
- Dikpati, M., de Toma, G. & Gilman, P. A. Predicting the strength of solar cycle 24 using a flux-transport dynamo-based tool. *Geophys. Res. Lett.* **33**, 5102 (2006).
- Choudhuri, A. R., Chatterjee, P. & Jiang, J. Predicting Solar Cycle 24 With a Solar Dynamo Model. *Phys. Rev. Lett.* **98**, 131103 (2007).
- Karak, B. B. & Nandy, D. Turbulent Pumping of Magnetic Flux Reduces Solar Cycle Memory and thus Impacts Predictability of the Sun's Activity. *Astrophys. J. Lett.* **761**, L13 (2012).
- Zharkova, V. V., Shepherd, S. J. & Zharkov, S. I. Principal component analysis of background and sunspot magnetic field variations during solar cycles 21–23. *Mon. Notices of RAS* **424**, 2943–2953 (2012).
- Hoeksema, J. T. *Structure and evolution of the large scale solar and heliospheric magnetic fields*. Ph.D. thesis, AA(Stanford Univ., CA.) (1984).
- Popova, E., Zharkova, V. & Zharkov, S. Probing latitudinal variations of the solar magnetic field in cycles 21–23 by Parker's Two-Layer Dynamo Model with meridional circulation. *Annales Geophysicae* **31**, 2023–2038 (2013).
- Zharkov, S., Zharkova, V. V. & Ipson, S. S. Statistical Properties Of Sunspots In 1996 2004: I. Detection, North South Asymmetry And Area Distribution. *Sol. Phys.* **228**, 377–397 (2005).

19. Shepherd, S. J., Zharkov, S. I. & Zharkova, V. V. Prediction of Solar Activity from Solar Background Magnetic Field Variations in Cycles 21–23. *Astrophys. J.* **795**, 46 (2014).
20. Clette, F., Svalgaard, L., Vaquero, J. M. & Cliver, E. W. Revisiting the Sunspot Number. A 400-Year Perspective on the Solar Cycle. *Space Sci. Rev.* **186**, 35–103 (2014).
21. Schmidt, M. & Lipson, H. Distilling free-form natural laws from experimental data. *Science* **324**, 81–85 (2009).
22. Kingmill, T. W. A 300-Year Climatic and Solar Cycle. *Nature* **73**, 413–414 (1906).
23. Wagner, G. *et al.* Presence of the Solar de Vries Cycle (~205 years) during the Last Ice Age. *Geophys. Res. Lett.* **28**, 303–306 (2001).
24. Zhao, J., Bogart, R. S., Kosovichev, A. G., Duvall, T. L., Jr. & Hartlep, T. Detection of Equatorward Meridional Flow and Evidence of Double-cell Meridional Circulation inside the Sun. *Astrophys. J. Lett.* **774**, L29 (2013).
25. Stix, M. Differential rotation and the solar dynamo. *Astron. and Astrophys.* **47**, 243–254 (1976).
26. Jouve, L. *et al.* A solar mean field dynamo benchmark. *Astron. and Astrophys.* **483**, 949–960 (2008).
27. Zeldovich, I. B., Ruzmaikin, A. A. & Sokolov, D. D. (eds.). *Magnetic fields in astrophysics* vol. 3 (1983).
28. Jolliffe, I. T. *Principal Component Analysis, 2nd edition*. (Berlin: Springer Series in Statistics, 2002).
29. Faber, N. M., Meinders, N. J., Geladi, P., Sjostrom, M., Buydens, L. M. C. & Kateman, G. Random error bias in principal component analysis. I: Derivation of theoretical predictions. *Analytica Chimica Acta* **304**, 257–271 (1995).
30. Wentzell, P. D. & Lohnes, M. T. Maximum likelihood principal component analysis with correlated measurement errors: theoretical and practical considerations. *Chemometrics and Intelligent Laboratory Systems* **45**, 65–85 (1999).
31. Schiesser, W. E. *The Numerical Method of Lines* 293 (Academic Press, New York, 1991).
32. Popova, E. P. & Potemina, K. A. Modeling of the solar activity double cycle using dynamical systems. *Geomagnetism and Aeronomy* **53**, 941–944 (2013).
33. Hathaway, D. H. Supergranules as Probes of the Sun's Meridional Circulation. *Astrophys. J.* **760**, 84 (2012).
34. Dikpati, M., Gilman, P. A., de Toma, G. & Ulrich, R. K. Impact of changes in the Sun's conveyor-belt on recent solar cycles. *Geophys. Res. Lett.* **37**, 14107 (2010).
35. Belucz, B., Dikpati, M. & Forgacs-Dajka, E. A Babcock-Leighton solar dynamo model with multi-cellular meridional circulation in advection- and diffusion-dominated regimes. *ArXiv e-prints* (2015).

Acknowledgements

The authors wish to thank the Directorate and staff of the Wilcox Solar Observatory (Stanford) for providing online the synoptic magnetic field data of the Sun for the past 4 solar cycles. VZ and SZ wish to acknowledge that this research was started during the EU Framework 5 grant 'European Grid of Solar Observations', grant IST-2001-32409.

Author Contributions

V.Z. and S.Z. conceived the experiment and analysed the data with PCA, S.S. conducted the data prediction with Euriqa, E.P. developed the model, V.Z. and E.P. analysed the results. All authors reviewed the manuscript.

Additional Information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Zharkova, V. V. *et al.* Heartbeat of the Sun from Principal Component Analysis and prediction of solar activity on a millennium timescale. *Sci. Rep.* **5**, 15689; doi: 10.1038/srep15689 (2015).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>