

Long-term solar activity and terrestrial connections.* Part I: theory

N. D. Diamantides

Department of Geography, Kent State University, P.O. Box 5190, Kent, Ohio 44242, USA

Received: 25 March 1996 / Revised: 16 August 1997 / Accepted: 5 September 1997

Abstract. The research task described herein aims at the structuring of an analytical tool that traces the time course of geophysical phenomena, regional or global, and compares it to the course of long-term solar conditions, long-term meaning decades or a few centuries. The model is based on the premise that since in a last analysis the preponderance of atmospheric, hydro-spheric, and, possibly, some aspects of geospheric phenomena are, or have been, powered by energy issuing from the sun – either now or in the past, the long-term behavior of such phenomena is ultimately “connected” to long-term changes occurring in the sun itself. Accordingly, the proposed research firstly derives and models a stable *surrogate pattern* for the long-term solar activity, secondly introduces a *transfer-function* algorithm for modeling the connection between the surrogate and terrestrial phenomena viewed as partners in the connection, and thirdly probes the connection outcome for *episodic* or *unanticipated effects* that may arise due to the fact that in the present context, the connection, should it exist, is very likely nonlinear. Part I of the study presents the theory of the concept, while Part II demonstrates the concept’s pertinence to a number of terrestrial phenomena.

Key words. Solar activity · Kolmogorov algorithm

1 Introduction

The particular aim of this essay is to examine the macrodynamic aspects of solar activity without invoking the particular agents that produce it (such as plasma

physics, helioseismology, planet-induced tidal effects, etc), and to do this in view of the presently incomplete knowledge of the actual mechanisms involved. Consequently, the outcome of this research will be descriptive of behavior morphology, and not necessarily explanatory of causation as far as the phenomena in question are concerned. In this context, let it be remembered that Isaac Newton is often cited in support of the school which prefers to establish measurements rather than to ascertain causes, “*hypotheses non fingo*”, as he said.

However, he was far from suggesting this as a limitation of scientific research in general, as the following passage from his *Opticks* attests:

To derive two or three Principles of Motion ...would be a very great step in Philosophy, though the Causes of those Principles were not yet discover’d: And therefore I scruple not to propose the Principles of Motion above-mention’d, they being of very general Extent, and leave their causes to be found out. (Samuel, 1952).

Our probing is based on the premise that since in a last analysis the preponderance of atmospheric, hydro-spheric, and possibly some aspects of geospheric phenomena are or have been powered by energy issuing from the sun, their long-term behavior reflects long-term changes occurring in the sun itself. Hence, the proposed research

1. proposes and models a stable *surrogate pattern* for the long-term solar activity;
2. introduces a *transfer-function* algorithm for modeling the connection between this surrogate and terrestrial phenomena for which solar activity has relevance;
3. probes the connection outcome for episodic or unanticipated events, given that the connection, should it exist, is very likely nonlinear.

2 Methodology

It must be stated from the beginning that the basic nature of the analysis attempted here is purposely cross-

* This essay has been the basis of two earlier papers, one presented at “The 8th International Conference of the European Union of Geosciences” in Strasbourg, France, 9–13 April 1995; and the other at “The 5th International Conference on Precipitation” in Elounda, Crete, Greece, 14–16 June 1995.

disciplinary, partaking of both physical science and engineering. This stems from the conviction that theoretical tools which sent man to the moon and are now preparing equipment for the unmanned exploration of Mars, can also assist geoscience in deciphering messages from the sun, should they exist. However, regarding solar-terrestrial connections, such geophysically basic occurrences as the El Niño in the Pacific, the temperature of either hemisphere, volcanos throughout the globe, and tree-rings in different locales, are occurrences which hardly interest most engineering analysts. By the same token, such potent tools of the latter as transfer-functions and system-response algorithms that deal with dynamic phenomena in general, do not people the tool kit of geophysicists at large. Against this background, our study takes the position that in order to proceed intelligently with regard to solar-terrestrial connections, the two types of skill need to complement each other on the task – difficulties notwithstanding (Keyfitz, 1992, 1994) – and that the Kolmogorov algorithm is the medium that can bind the two together in the present context: its principal worth being that it implicitly captures the dynamics propelling a given phenomenon from the phenomenon's past course, even where no adequate theoretical knowledge of such dynamics presently exists. The algorithm's detailed derivation in Part I is for the benefit of the geoscience element of the interdisciplinary team.

The analysis begins with the observation that solar activity is not only in tune with the familiar 11-year cycle of the sunspot number, but, in addition, varies along a time-scale measured in centuries, as the Sporer and Maunder Minima demonstrate (Eddy, 1976; Hughes, 1977). One approach to examining the activity's long-term behavior will be to consider the sunspot number as its surrogate and, through it, structure a mathematical model of solar activity potentially useful to the study of whatever phenomena such activity may have a bearing upon.

The suggested method proceeds with the task of modeling the sunspot number in the long run based on the following tenet and its corollary:

The Constrain Tenet. Any solar process involving the exchange or release of energy $y(t)$, has an upper bound Y , and a time rate of change dy/dt that increases with both the current value of $y(t)$ and the remaining potential $Y - y(t)$ for further change. This renders both y and dy/dt bounded. The modality is not new, having been introduced by P. F. Verhulst in 1844 in connection with population growth (Lotka, 1956).

The Corollary. Since the constrain tenet means that the process $y(t)$ is nonlinear, the process cannot only give rise to a basic oscillation f_0 , but also cause:

1. the appearance of new incommensurable frequencies f_1, f_2, \dots ,
2. the entrainment of these frequencies so that (for example) f_1, f_2 approach values f_0 ,
3. the sudden appearance of aperiodic spectra,
4. the appearance of subharmonics $f_1/2, f_2/4$ successively (McDonald, 1976).

Put in terms of a differential equation, the tenet requires that the process evolves as:

$$\frac{dy}{dt} = cy(Y - y) , \quad (1)$$

where $c = c(t)$ is a factor, facilitative or inhibitory, established by the determinants of the process in question, which therefore may be called the *mediating factor*. The expression in Eq. (1) is a Bernoulli-type differential equation when c is time dependent, and becomes the well-known logistic law when c is constant. The equation's general solution is:

$$y(t) = \frac{1}{(1/y_{\min} - 1/y_{\max}) \exp(-y_{\max} \int c(t) dt) + 1/y_{\max}} ,$$

and

$$\int c(t) dt = \ln \frac{1/y_{\min} - 1/y_{\max}}{1/y(t) - 1/y_{\max}} . \quad (2)$$

Depending on $c(t)$, this basic expression can model a great variety of physical (as well as societal) phenomena and, being nonlinear, can lead to outcomes of unexpected complexity, including cases of the so-called deterministic chaos (May, 1976; Modis and Bebecker, 1992; Mullin and Price, 1989). The basic features of Eq. (1) are that:

1. $dy/dt = 0$ when $y = 0$ or $y = Y$,
2. $dy/dt = \max$ when $2y = Y - 1/c^2 dc/dt$.

To compute the two extremes y_{\min} and y_{\max} for Eq. (2) from the data of a particular case, we may use the expedient of approximating either end y_{end} by an exponential curve; it is easy to show that three equidistant points of such a curve will produce:

$$y_{\min} \text{ OR } y_{\max} = \frac{y_{\text{end}}(t)^2 - y_{\text{end}}(t-a)y_{\text{end}}(t+a)}{2y_{\text{end}}(t) - y_{\text{end}}(t-a) - y_{\text{end}}(t+a)} . \quad (4)$$

The two features in Eq. (3) imply a cusp-like shape for dy/dt , simple when $c(t) = c$ (constant), and more complicated when $c(t)$ varies in time. When, as in the present context, the mediating factor is a priori unknown, the morphology of dy/dt may be approximated by the expression:

$$\frac{dy}{dt} = \exp(at^2 + bt + k), \quad a < 0, \quad a, b, k = \text{const.} \quad (5)$$

as an intermediate step. This simple function is chosen because it allows for a single peak only and, therefore, for the capture of the very core of the growth process. The condition upon the parameter a ensures that in the limit dy/dt settles at a finite nonnegative level, as it should for the constrain tenet to be satisfied. Given that the root of the derivative of dy/dt in Eq. (5) defines the time t_p of the dy/dt peak as $t_p = -b/2a$, it follows that when t_p is known, Eq. (5) becomes:

$$\frac{dy}{dt} = \exp[at(t - 2t_p) + k] . \quad (6)$$

Since the parameters a , k in Eq. (6) can be found from the data, $y(t)$ and Y are computable using Eq. (1) and, through them, the mediating factor $c(t)$.

3 Model synthesis

The long-term behavior of solar activity can be brought into focus by filtering out the ubiquitous annual mean relative Wolf sunspot number and its 11-year periodicity.

The simplest means of effecting this is through the 11-year moving average (Eddy, 1976), for which the symbol $z_{11}(t)$ will be used:

$$z_{11}(t) = \frac{1}{11} \sum_{t-10}^0 z(t) , \quad (7)$$

the time history of which may be seen in Fig. 1.

The $z(t)$ data on which the present analysis is based comes from McKinnon (1987) for the years 1700–1985, and from the monthly issues of *Sky and telescope* after 1985. It should be remarked that the reliability of the data deteriorates as one goes back in time (Eddy and Stuiver, 1977), but they are deemed accurate enough at least from 1800 to the present: the assessment is in agreement with the results of a study by Hoyt *et al.* (1994) who used 340 000 daily sunspot group counts from more than 350 observers active between 1610 and 1993.

Three prominent features characterize $z_{11}(t)$ over the course of the almost three hundred years for which data exist:

1. a continuously rising secular trend $T(t)$,
2. long-term surging $S(t)$,
3. a quasi-patterned transient $Tr(t)$.

Among them, $T(t)$ and $S(t)$ constitute the steady-state part of $z_{11}(t)$. Of these two, the $S(t)$ part gives a strong indication of the ‘*Gleissberg Cycle*’ (Eddy, 1976), represented by approximately two-and-a-half surges in

the spans 1705–1811, 1812–1905, and 1906–?. Quite obviously, the course of each surge is of the type implied by Eq. (1). All three surges ride over what may be called a *base-line* which parallels the secular trend, and all three feature a definite *bifurcation* in the neighborhood of the surge peak. In addition, a higher-frequency transient $Tr(t)$ is obvious in all three surge cases, causing a pair of subpeaks for each surge.

3.1 Secular trend $T(t)$

While the upward trend of $z_{11}(t)$ over the last three centuries is quite obvious in Fig. 1, its increase with time cannot be open ended if the constrain tenet is to be satisfied. However, since no peaking of $T(t)$ has been discerned over the length of the record, the analysis has to assume such a peaking at some future time t_p and, based on it, determine the course of $T(t)$ using Eq. (6) in the form:

$$\ln \frac{dT(t)}{dt} = at(t - 2t_p) . \quad (8)$$

The t_p value can be found by applying linear regression iteratively to the 1700–1991 record and selecting the t_p that minimizes the goodness-of-fit coefficient R^2 . The result, including year of the peak t_p , peak value T_p , and duration t_d , is given by:

$$\begin{aligned} a &= -0.410 \times 10^{-6} \text{ (} 0.031 \times 10^{-6} \text{)}, \\ k &= 3.41, \\ t_p &= 2100 \text{ AD,} \\ T_p &= 60.8, \\ t_d &= 800 \text{ years .} \end{aligned} \quad (9)$$

If, in violation of the constrain-tenet, the foregoing parameters were computed without imposing a peaking requirement, the result would be:

$$\begin{aligned} a &= 0.800 \times 10^{-5} \text{ (} 0.064 \times 10^{-5} \text{)}, \\ b &= -0.383 \times 10^{-4} \text{ (} 0.044 \times 10^{-5} \text{)}, \\ k &= 3.574, \\ t_p &= -b/2a = 2.4 \text{ years .} \end{aligned} \quad (10)$$

This implies unbounded growth for $T(t)$, due to the fact that the end of the data record used is the ascending part of the third surge presently at its peak. It is of interest, however, that the curve of Eq. (10) bottoms out in the year 2.4, or 1707 AD, which is the ending era of the *Maunder Minimum*.

For the base line $B(t)$, on the basis of which the surges are to be computed, the following simple definition was chosen:

$$\begin{aligned} B(t) &= T(t) - B \\ B &= 24.1 . \end{aligned} \quad (11)$$

This is simply the bounded curve of Eq. (9) lowered by k units of $z_{11}(t)$ to between the 1815 and 1906 minima of $z_{11}(t)$.

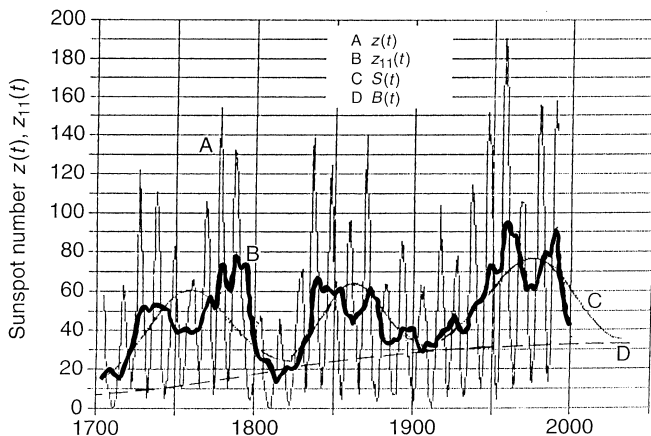


Fig. 1. Sunspot short-term and long-term

Table 1. Solar activity periods

name of maximum	timing	time between maxima (years)
Homeric	1375–700 BC	675
Greek	700–375 BC	325
Roman	375 BC–690 AD	1065
Medieval	690–1490 AD	800
Columbian	1490–1675 AD	185
Modern	1675–2475 AD	800 (?)

The duration of $T(t)$, and therefore $B(t)$, was established by first finding the base line’s value $B(1715) = 6.64$ at the end of the Maunder Minimum, and computing the future year during which the descending branch of $B(t)$ reaches the same level. To place the result in historical perspective, we observe that the contemporary trend $T(t)$ corresponds to the latest of Eddy’s historical ‘maxima’ of the sunspot activity’s envelope using ^{14}C as a guide (Stuiver and Quay, 1980). Of these, the last six have occurred in historical times as seen in Table 1, ‘‘Columbian’’ being our designation for the spurt of envelope activity that took place between the Sporer and Maunder Minima. Duration comparison between the two tabulations shows the ongoing Modern Maximum to be on par with the Medieval Maximum.

3.2 The surges $S(t)$

The morphology of the surging part of $z_{11}(t)$ may be looked into either collectively or individually.

3.2.1 Collective surging. From a collective viewpoint, particularly dominant over the post-Maunder era appear to be two oscillations with periods of about 179 and 110 years (Williams, 1981). The first oscillation in particular has attracted researchers’ attention due to the fact that it approaches the behavior of orbital phenomena in the solar system: Jose (1965) has shown this periodicity to be prominent in the motion of the sun about the center of mass of the solar system, and Gribbin and Plagemann (1974) point out that 179 years is the period of total planetary alignment during which every planet is in conjunction with every other planet on the same side of the sun. Wood (1972), in analyzing the connection between sunspot number and the height of the planetary gravity tide on the sun’s surface, observes that ‘‘the curve appears to repeat after about 170 or 180 years’’. Cohen and Lintz (1974), on the other hand, suggest that ‘‘although the 179-year periodicity does exist in the sunspot cycle, it is not caused by a long-term planetary excitation but is, instead, a beat phenomenon’’. By subjecting the 1844–1971 $z(t)$ data to a maximum-entropy spectral analysis, they found three dominant frequencies in the spectrum: 110, 10.9, and 9.7 years. Of these the latter two ‘‘interfere to produce

beats of roughly 187 years’’. In our own work we found that the 178.7- and 110-year components do make for a good representation of the long-term surging in $z_{11}(t)$. To arrive at this, the $S(t)$ were separated from $z_{11}(t) - B(t)$ and Eq. (8) was applied for each surge separately, leading to the cusps seen in Fig. 1, and the following parameters:

surge	S_I	S_{II}	S_{III}
a	–0.0007 (0.00005)	–0.0014 (0.00010)	–0.0006 (0.00004)
b	0.0631	0.1349	0.0786
k	2.497	0.488	1.238
R^2	0.565	0.655	0.820
area S_A	3435	1890	3370
peak	62.1	62.9	78.1
peak year	1758	1865	1973
duration	1705–1821	1821–1906	1906–2045.

(12)

In a follow-up step, the surge function $S_j, j = I, II, III$, were expressed in terms of the aforesaid sinusoids:

$$S_j(t) = \sum_n \sum_N \left[a_n \sin \left[\frac{2\pi}{N} (t - 1705) \right] + b_n \cos \left[\frac{2\pi}{N} (t - 1705) \right] \right] + k_0 \quad (13)$$

$n = 1, 2, \quad N = 178.7, 110, \quad 1705 \leq t \leq 1991,$

with the following results:

$$\begin{aligned} a_1 &= 4.423 (0.313), \\ b_1 &= -4.559 (0.322), \\ a_2 &= 6.334 (0.323), \\ b_2 &= -17.320 (0.315), \\ k_0 &= 25.35, \\ R^2 &= 0.93. \end{aligned} \quad (14)$$

Projecting into the future, this model suggests that the core (not counting the transient) of the next surge S_{IV} will peak in the year 2090 at the level of 43.9 units above the base line, as Fig. 2 portrays it. This renders the ongoing surge S_{III} the strongest in the span 1700–2100 of four centuries, with a core peak of 48.3 units above the base line. With the base line included, S_{III} and S_{IV} emerge about equal (79.8 units).

3.2.2 Individual surges. While the use of the Eq. (6) expedient serves a purpose in that it gives an overview of the long-term activity, it is not conducive to probing the morphology of the surges’ internal structure. For this, not only the core part $S(t)$, but the entire surge $S(t) + Tr(t)$ needs to be dealt with, i.e.:

$$S(t) + Tr(t) = z_{11}(t) - B(t). \quad (15)$$

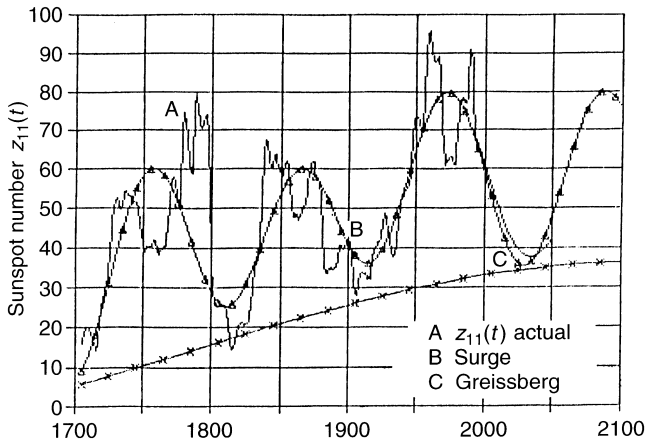


Fig. 2. Collective surging

After finding the total area S_A under the curve of the right-hand member, use of it and of $s(t)$ in Eq. (1) produces the mediating factor $c(t)$, that is the key structural element of the model, and the nature of which can be seen in Fig. 3. This is manifestly a trigger type of occurrence whose time behavior is made up of the combination of a growth term and a decay term operating concurrently,

$$c_0(t) = \begin{cases} c_0 & \text{if } t = 0 \\ c_1 e^{+\alpha t} + c_2 e^{-\beta t}, & \text{if } t > 0 \end{cases} \quad (16)$$

with the parameter values of each case seen in the following equation and the process pictured in Fig. 4.

surge	S_I	S_{II}	S_{III}
c_1	0.362×10^{-4}	1.553×10^{-4}	0.351×10^{-4}
c_2	0.032×10^{-5}	0.743×10^{-5}	0.104×10^{-5}
α	0.055 (0.004)	0.022 (0.002)	0.029 (0.002)
β	-0.029 (-0.002)	-0.041 (0.003)	-0.215 (0.016).

(17)

Of interest is the fact that the observed dip at each $S(t)$ peak occurs near the crossing of the two $c_0(t)$ branches in Eq. (16). The term c_1 depicts a positive feedback process, which at some level triggers the negative

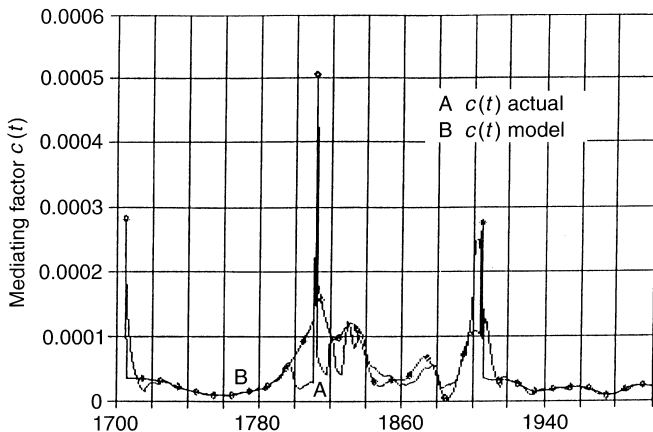


Fig. 3. Mediating factor (actual and model)

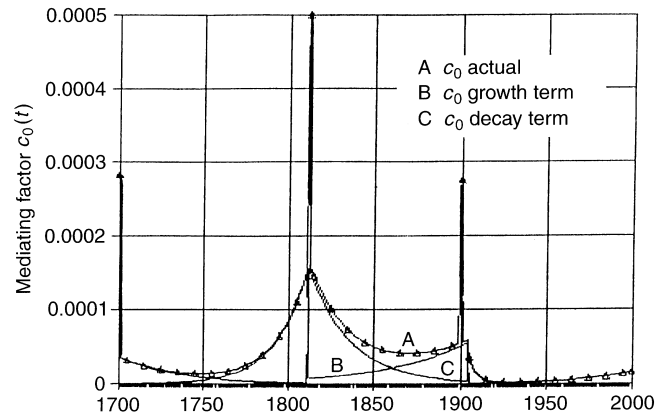


Fig. 4. Mediating factor (growth and decay)

feedback process represented by the negative feedback term c_2 of the next surge's mediating factor. Interestingly enough, this modality replicates a relationship between solar activity and the earth's magnetic field: the diurnal variation of the horizontal component of this field on abnormally quiet days (AQD) at the time of sunspot minimum consistently anticipates the magnitude of the subsequent sunspot minimum (Brown, 1974). This pattern implies "that the size of solar activity maximum is determined at the very beginning of a cycle, or perhaps the very end of the preceding cycle from the 'depth' of the solar minimum" (Brown, 1974).

With $c_0(t)$ known, integration of the expression in Eq. (1) produces the equation $s(t)$ of each individual surge as:

$$s(t) = \frac{AS_A c_0(t) \exp(-J)}{[A \exp(-J) + 1]^2}, \quad (18)$$

$$A = \frac{1}{s_0} - \frac{1}{S_A},$$

$$J = S_A \left[\frac{c_1}{\alpha} (e^{+\alpha t} - 1) - \frac{c_2}{\beta} (e^{-\beta t} - 1) \right] + i_0,$$

where s_0 is the value of $S(t)$ at the time t_0 of the surge onset, and i_0 the integration constant. The values of these two parameters, along with that of the model S_A , are given as

surge	S_I	S_{II}	S_{III}
S_A	3560	1990	3630
S_0	9.82	3.87	2.27
i_0	0.002	0.003	0.002.

(19)

Table 2 lists the projected values of the steady-state part of $z_{11}(t)$ in column A, from the present to the end of surge S_{III} . In column B of the table an alternate projection is listed, based on the observation that the surge S_{III} appears to be symmetric about the 1977 position and, therefore:

$$S_{III}(1977 + t) = S_{III}(1977 - t) + k, \quad k = -0.95, \quad t \geq 14. \quad (20)$$

Table 2. $z_{11}(t)$ projection

year	$z_{11}(t)$		year	$z_{11}(t)$	
	A	B		A	B
1990	94.5		2007	57.9	58.6
1991	93.5	94.6	2008	56.4	56.3
1992	91.8	89.9	2009	54.9	54.1
1993	89.5	92.6	2010	53.5	51.9
1994	86.8	94.6	2011	52.1	49.9
1995	83.9	92.6	2012	50.8	46.4
1996	80.9	89.6	2013	49.9	44.9
1997	78.0	80.7	2014	48.3	43.7
1998	75.2	70.8	2015	47.2	42.8
1999	72.6	68.2	2016	46.2	42.1
2000	70.2	69.3	2017	45.3	41.6
2001	68.0	72.9	2018	44.5	41.3
2002	66.0	70.5	2019	43.8	41.2
2003	64.2	68.1	2020	43.2	41.1
2004	62.5	65.7	2025	40.7	40.9
2005	60.9	63.3	2030	38.4	37.5
2006	59.4	61.0	2035	36.6	32.8

The restriction upon t places the projection at the year 1992 or later, while the constant k makes the average value of the difference between columns A and B equal to zero. It is curve B that will be used in Part II of this study.

The following surge characteristics are presented at this point in order to effect comparisons:

	peak 1	trough	peak 2	surge
	position	position	position	duration
S_I	0.301	0.528	0.783	1 (= 106 years)
S_{II}	0.293	0.521	0.783	1 (= 92 years)
S_{III}	0.400	0.540	0.637	1 (= 125 years).

(21)

3.3 The transient $Tr(t)$

The relatively shorter period undulations $Tr(t)$ exhibited by $z_{11}(t)$, in addition to the long-span surges, are accounted for by the basic tenet's corollary and therefore have to be represented by a term $c_r(t)$ in the mediating factor $c(t)$. Accordingly, and for each surge separately, the residual:

$$c_r(t) = c(t) - c_0(t) \tag{22}$$

was regressed on the sum of the first three subharmonics of the 11-year cycle and of the Gleissberg cycle (8th subharmonic) mentioned earlier:

$$c_r(t) = \sum_N \left[a_N \sin\left(\frac{2\pi}{11.1N}t\right) + b_N \cos\left(\frac{2\pi}{11.1N}t\right) \right] + k, \tag{23}$$

$N = 2, 3, 4, 8 .$

The reason for choosing this set of N -values is that the first three subharmonics have been found, via Fourier analysis, in at least two widely separated terrestrial occurrences in which the solar factor is found implicated

(Clark, 1988; Williams, et al. 1985). The eighth subharmonic was added heuristically, because its inclusion was found to triple the value of the regression's goodness-of-fit coefficient R^2 . This subharmonic is, of course, the Gleissberg cycle already mentioned.

The a_N, b_N values found through regression are shown to be:

coef. $\times 10^6$	S_I	S_{II}	S_{III}
a_2	0.397 (0.043)	-7.348 (0.033)	-1.766 (0.049)
b_2	0.441 (0.074)	14.815 (0.088)	0.385 (0.029)
a_3	0.897 (0.061)	-24.096 (0.077)	-1.294 (0.059)
b_3	0.506 (0.055)	-0.197 (0.068)	-1.817 (0.070)
a_4	1.732 (0.071)	-8.602 (0.081)	2.808 (0.091)
b_4	-2.384 (0.067)	5.618 (0.077)	2.843 (0.087)
a_8	7.850 (0.059)	4.271 (0.090)	-0.373 (0.092)
b_8	0.601 (0.078)	11.552 (0.045)	-0.084 (0.009)
k	0.136	3.093	0.855
R^2	0.902	0.789	0.758.

(24)

Figure 5 shows the complete model $\bar{z}_{11}(t)$. It is worth noticing the similarity between $\bar{z}_{11}(t)$ and J.A. Eddy's annual mean sunspot numbers at maxima in the 11-year cycle (Eddy, 1976, Fig. 8). The degree of agreement between actual and modeled surge may be judged based upon the mean and standard deviation of the two:

	S_I		S_{II}		S_{III}	
	actual	model	actual	model	actual	model
avg	32.33	32.51	21.26	21.29	32.21	32.98
std	16.60	16.66	15.01	14.40	18.94	20.90.

(25)

Before leaving the topic of the long-term $z_{11}(t)$ surging of solar activity, we may point out the fact that the trigger-like nature of the "mediating factor" is akin to

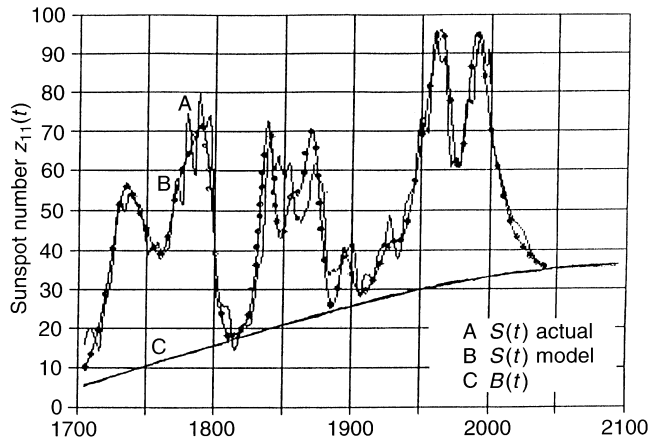


Fig. 5. Surges (actual and model)

the “eruption mechanism” that has been postulated for the 11-year sunspot cycle $z(t)$ itself (Kiepenheuer, 1959). As Dicke puts it, “It has become clear by now that the rise and fall in the number of spots is due to a number of practically independent individual processes: thus the idea of a true periodic phenomenon was dropped in favor of the so-called “eruption hypothesis” (Dicke, 1978).

3.4 The derivative $Dz_{11}(t)$

As in most dynamic processes, the time-derivative $Dz_{11}(t)$ of $z_{11}(t)$ can be instrumental in cases where $z_{11}(t)$ is a casual factor for the process at hand. It can be computed either from Eq. (1) or from the first difference of $z_{11}(t)$ as a surrogate which, according to Eq. (7), results in:

$$z(t) = 11Dz_{11}(t) + z(t - 11) , \tag{26}$$

and offers the means of projecting the sunspot number 11 years ahead.

Since Eq. (18) makes possible the extension of the model $z_{11}(t)$ beyond the present, a projection is possible with an accuracy commensurate with the fidelity of the

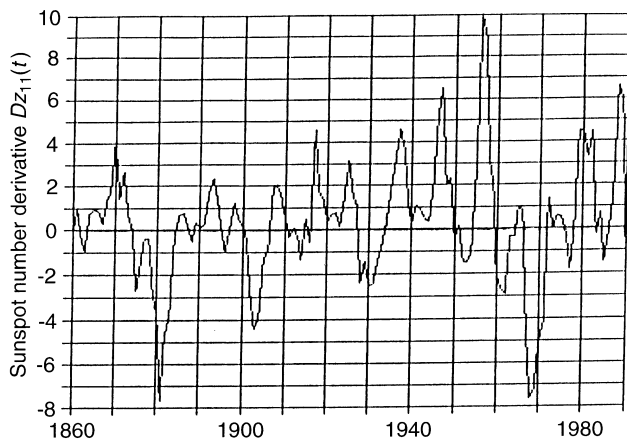


Fig. 6. Long-term sunspot derivative

model $z_{11}(t)$; Fig. 6 shows the latter, revealing a regularity, which we discuss further later in this study.

4 Model implementation

The analytical means through which the present model is to be applied in seeking possible links between long-term solar activity on the one hand, and terrestrial phenomena on the other, range from the simple to the fairly sophisticated. In the sense that this means “connecting” one function of time $f(t)$ with another $g(t)$, the simplest relationship will be the direct one:

$$g(t) = af(t) + b \tag{27}$$

$a, b = \text{const.} ;$

or, when the data ranges warrant it,

$$g_N(t) = \frac{g(t) - \bar{g}}{\sigma_g} = a \frac{f(t) - \bar{f}}{\sigma_f} + b , \tag{28}$$

where \bar{g} and \bar{f} are the averages, σ_g and σ_f the standard deviations of the two functions, and $g_N(t), f_N(t)$ the *normalized* $g(t)$ and $f(t)$. This simple transformation brings the two data ranges within common bounds, and when necessary, can be rendered more effective if before normalization, $g(t)$ were “stretched” to:

$$g_\infty(t) = g(t)^m . \tag{29}$$

Normalization allows for a meaningful quantitative comparison of the contribution of multiple inputs $f_1(t) \dots f_n(t)$ affecting a common output $g(t)$. Normalization, therefore, will be applied routinely in the entries of Part II of the study.

Where a more complex dynamic input-output connection is involved or suspected than that of Eq. (27), the powerful Kolmogorov method of signal-processing and *system-response* may be applied to advantage (Kolmogorov, 1941). Signal-processing pertains to either the interpolation or extrapolation of $f(t)$, while system-response addresses underlying causality or co-evolution; generally, both cases have to cope with contamination from extraneous noise. The method has proved most effective in signal detection and estimation, as well as in control system analysis and design. A simplified derivation of the method by this author is given in the following.

Kolmogorov’s formulation of the signal extrapolation case assumes that a function $f(t)$ of t , where both t and $f(t)$ are real variables, is given in terms of equally spaced samples

$$f(-n), f(-n + 1), \dots, f(-1), f(0) , \tag{30}$$

$f(0)$ being the current value, and that its value $f(k)$, k units of time into the future, is to be predicted. It is assumed that $f(t)$ is weakly stationary and, therefore,

1. that it has a finite second moment

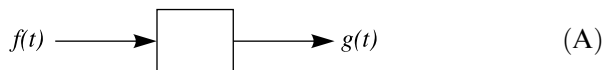
$$E[f(t)^2] < \infty , \tag{31}$$

Table 3. Autocorrelation functions (note: the correlations Q_{ff} are for the functions $f(t) - f_{av}$, where $[z]_{av} = 59.25$, $[z_{11}]_{av} = 48.61$, $[Dz_{11}]_{av} = 2.97$; the time-spans are 1890–1990 for $z(t)$, 1755–1990 for $z_{11}(t)$, and 1740–1990 for $Dz_{11}(t)$)

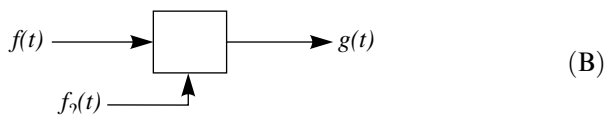
year	$Q_{ff}(\cdot)$			year	$Q_{ff}(\cdot)$		
	$z(t)$	$z_{11}(t)$	$Dz_{11}(t)$		$z(t)$	$z_{11}(t)$	$Dz_{11}(t)$
0	1.000	1.000	1.000	26	-0.346	-0.032	-0.071
1	0.806	0.974	0.741	27	-0.375	-0.065	-0.126
2	0.404	0.930	0.358	28	-0.256	-0.095	-0.149
3	-0.035	0.879	0.026	29	-0.032	-0.121	-0.148
4	-0.355	0.831	-0.135	30	0.231	-0.145	-0.108
5	-0.495	0.787	-0.146	31	0.423	-0.167	-0.067
6	-0.428	0.747	-0.035	32	0.489	-0.183	-0.016
7	-0.176	0.707	0.129	33	0.381	-0.209	-0.011
8	0.181	0.663	0.262	34	0.158	-0.228	-0.036
9	0.528	0.612	0.264	35	-0.078	-0.241	-0.078
10	0.735	0.555	0.080	36	-0.241	-0.247	
11	0.706	0.498	-0.191	37	-0.297	-0.246	
12	0.471	0.448	-0.270	38	-0.254	-0.243	
13	0.145	0.407	-0.238	39	-0.116	-0.236	
14	-0.147	0.373	-0.117	40	0.053	-0.230	
15	-0.343	0.342	0.016	41	0.202	-0.224	
16	-0.400	0.309	0.109	42	0.262	-0.221	
17	-0.318	0.274	0.097	43	0.230	-0.220	
18	-0.105	0.236	0.014	44	0.127	-0.218	
19	0.187	0.197	-0.084	45	0.009	-0.213	
20	0.450	0.162	-0.186	46	-0.102	-0.202	
21	0.602	0.130	-0.209	47	-0.178	-0.187	
22	0.581	0.101	-0.144	48	-0.224	-0.171	
23	0.392	0.071	-0.053	49	-0.211	-0.157	
24	0.096	0.039	0.001	50	-0.141	-0.147	
25	-0.172	0.004	-0.001				

One feature of the Kolmogorov algorithm as presented here is that its definition in Eqs. (33) and (39) does not provide for a constant term, having been developed originally in a communication-signal context. Care, therefore, needs to be exercised when applied to signals from other realms.

Equally challenging in the context of the present theme is the pertinence of the cross-correlation function $Qfg(\cdot)$: it has to do with the case where the cause side of the presumed cause/effect relationship $g(t) =$ function of $f(t)$ per diagram A:



is actually more complex as in diagram B,



but the factor $f_2(t)$ has been omitted for lack of knowledge regarding the innards of the phenomenon at hand. In this circumstance application of the Kolmogorov algorithm on the system A instead of B will leave a residual in which the presence of $f_2(t)$ remains to be detected and extracted. This of course can be a difficult task, especially when incommensurable frequencies enter the picture. A case in point is terrestrial climate where not only the energy input by the sun is a factor, but contaminants, either manmade (Plantico *et al.*, 1990) or subterranean (Schneider and Mass, 1975;

Lamb, 1970), which are released into the atmosphere play a significant role. Only in the case of known factors $f(t), f_2(t)$ acting additively can the algorithm in Eq. (45) be changed explicitly to represent the case.

For the record, Table 4 gives the actual and model values of the long-term solar activity's surrogate $z_{11}(t)$.

5 Sunspot number projection

We conclude Part I of this essay by demonstrating the model's ability to effect an 11-year projection of the annual sunspot number itself over the past 150 years; this via Eq. (33).

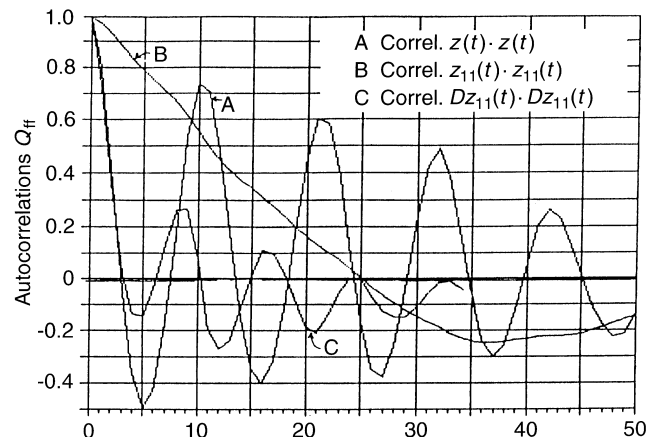


Fig. 7. Component autocovariances

Table 4. Long-term solar activity $z_{11}(t)$

year	z_{11}	year	z_{11}	year	z_{11}	year	z_{11}	year	z_{11}
1790	72.94	1830	31.01	1870	57.07	1910	33.99	1950	73.1
1791	71.28	1831	33.94	1871	58.47	1911	33.65	1951	73.25
1792	70.55	1832	35.84	1872	61.15	1912	33.73	1952	71.79
1793	71.31	1833	36.25	1873	61.79	1913	33.4	1953	70.27
1794	72.96	1834	37.28	1874	61.85	1914	32.05	1954	69.19
1795	73.97	1835	41.68	1875	59.13	1915	32.55	1955	71.77
1796	73.24	1836	51.22	1876	57.38	1916	31.96	1956	81.64
1797	66.28	1837	60.49	1877	57.03	1917	36.52	1957	90.51
1798	54.65	1838	65.27	1878	56.67	1918	38.21	1958	93.53
1799	43.37	1839	67.23	1879	53.8	1919	39.58	1959	95.59
1800	33.95	1840	67.01	1880	50.01	1920	39.92	1960	93.55
1801	28.87	1841	63.9	1881	42.31	1921	40.6	1961	90.83
1802	26.91	1842	61.75	1882	37.62	1922	41.37	1962	87.93
1803	25.37	1843	60.23	1883	33.72	1923	41.57	1963	87.6
1804	25.43	1844	60.82	1884	33.47	1924	42.96	1964	87.26
1805	25.54	1845	63.26	1885	34.15	1925	46.12	1965	88.24
1806	26.15	1846	63.68	1886	34.92	1926	47.62	1966	89.05
1807	25.62	1847	61.59	1887	35.08	1927	48.7	1967	84.7
1808	25.81	1848	60.35	1888	34.57	1928	46.33	1968	77.04
1809	25.66	1849	59.82	1889	34.93	1929	44.9	1969	69.83
1810	25.05	1850	58.08	1890	35.03	1930	42.36	1970	64.87
1811	23.86	1851	58.07	1891	35.33	1931	39.96	1971	60.72
1812	21.23	1852	59.65	1892	37.03	1932	38.6	1972	62.08
1813	18.25	1853	61	1893	39.35	1933	37.83	1973	62.13
1814	15.59	1854	61.9	1894	40.65	1934	38.09	1974	62.73
1815	14.49	1855	61.15	1895	40.69	1935	39.85	1975	63.21
1816	14.82	1856	57.89	1896	39.75	1936	43.07	1976	62.98
1817	16	1857	54.36	1897	39.82	1937	47.66	1977	61.21
1818	17.82	1858	50.39	1898	41.05	1938	51.35	1978	61.09
1819	19.22	1859	47.58	1899	41.54	1939	52.35	1979	65.6
1820	20.41	1860	47.54	1900	41.74	1940	52.62	1980	70.07
1821	21	1861	48.5	1901	41.34	1941	53.69	1981	73.35
1822	21.24	1862	48.01	1902	38.55	1942	54.55	1982	77.83
1823	20.95	1863	47.09	1903	34.14	1943	55.02	1983	77.62
1824	20.61	1864	47.82	1904	30.22	1944	55.37	1984	78.34
1825	20.85	1865	48.72	1905	28.9	1945	57.6	1985	76.83
1826	20.94	1866	49.59	1906	27.97	1946	62.74	1986	76.64
1827	21.28	1867	49.86	1907	29.81	1947	69.27	1987	78.15
1828	23.38	1868	51.22	1908	31.84	1948	71.26	1988	84.75
1829	26.74	1869	52.96	1909	33.4	1949	73.55	1989	90.68
								1990	89.5

A reliable prediction of the 11-year cycle “could be of great importance to many fields of human activity; telecommunications, space exploration, and climate forecasting being among them” (Fyodorof *et al.*, 1996). Over the years the list of attempts to produce a solution has been long: see for instance, Wilson (1985), McNish and Lincoln (1987), Withbroe (1989), and Friis-Christensen and Lassen (1991). Although the descriptive schemes have been many, and the numerical models range from simple to fairly elaborate, none has achieved the goal, including those based on a Fourier transform. The disappointment is due, of course, to the limited knowledge of the dynamics of the astrophysical process taking place. It is the view of this study that the Kolmogorov algorithm can help with the quest, and produce a better descriptor. For this, the suggested modus operandi proceeds in three successive steps:

1. Generate an 11-year Kolmogorov projection $z_a(t)$ of $z(t)$ based on three past equidistant states of it: $z(t - 33)$, $z(t - 22)$, and $z(t - 11)$.

2. Search the residual $r(t) = z(t) - z_a(t)$ for the presence of both the basic 11-year solar cycle and the long-term $z_{11}(t)$ or its derivative .
3. Finally, apply regression analysis to the residual $r(t)$ as a function of the terms found in Step 2.

The first step requires solving the system:

$$\begin{vmatrix} Q(0) & Q(11) & Q(22) \\ Q(11) & Q(0) & Q(11) \\ Q(22) & Q(11) & Q(0) \end{vmatrix} = \begin{vmatrix} a_{11} \\ a_{22} \\ a_{33} \end{vmatrix} \begin{vmatrix} Q(11) \\ Q(22) \\ Q(33) \end{vmatrix}, \quad (53)$$

from which, based on Table 3 for the values of $Q(\cdot)$ of the autocorrelation of $z(t)$, the following transfer function is computed:

$$\begin{matrix} i & W(i) \\ -33 & -0.00179 \\ -22 & +0.25351 \\ -11 & +0.97182, \end{matrix} \quad (54)$$

effecting the projection:

$$z_a(t) = \sum_i z(t - i) . \tag{55}$$

Since $z_a(t)$ and $z(t)$ have a concurrent part, the former can be improved by making it:

$$z_a(t) = mz(t) + n \tag{56}$$

and finding the parameters a, b through linear regression:

$$\begin{aligned} m &= 0.855 (0.067), \\ n &= 15.270, \\ R^2 &= 0.532 . \end{aligned} \tag{57}$$

This step serves to compensate for the absence of a constant term in the definition of the Kolmogorov algorithm mentioned earlier.

In the residual $r(t)$ of the operation described by Eq. (56), the presence of Dz_{11} is discerned, and in the same residual's autocorrelation function (Fig. 8) the presence of the sunspot number's basic 11-year cycle is obvious. Hence, $r(t)$ is set as:

$$r(t) = a \sin(2\pi/11 * t) + b \cos(2\pi/11 * t) + cDz_{11}(t) + d , \tag{58}$$

from which a new regression produces:

$$\begin{aligned} a &= -9.957 (0.755), \\ b &= -0.867 (0.755), \\ c &= 9.802 (0.186), \\ d &= -2.326, \\ R^2 &= 0.949, \\ t_0 &= 1830 . \end{aligned} \tag{59}$$

Thus, the complete 11-year projection expression becomes:

$$z(t + 11) = z_a(t) + r(t + 11) . \tag{60}$$

In this expression the $Dz_{11}(t + 11)$ term present in $r(t + 11)$ comes from the actual $z_{11}(t)$ up to the year 1995 and the model $z_{11}(t)$ beyond that.

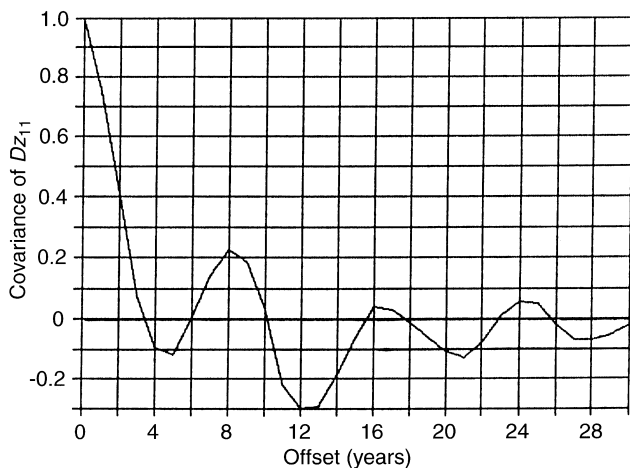


Fig. 8. Autocovariance of long-term derivative

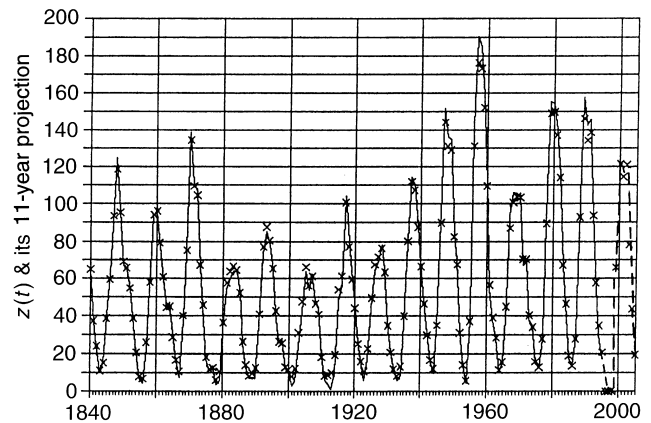


Fig. 9. Sunspot number 11-year projection

The final outcome may be seen in Fig. 9, while Table 5 lists the 1990–2005 projections. The strength of the approach can be assessed by the fact that in the 1835–1995 record of fifteen $z(t)$ maxima and fourteen minima, the model missed only the 1968 maximum by 1 year and the 1913 minimum by 2 years, which is smaller than the 2.3 years suggested by Fyodorov *et al.* (1996) as the limit for useful prediction.

Equally important is the fact that over the same time-span the Kolmogorov term of the 11-year prediction captured 86% of the $z(t)$ signal; this is based on the ratio of the standard deviation (37.6) of $z(t)$ -predicted versus (41.5) of $z(t)$ -actual. For the complete prediction algorithm the ratio is 0.94. As far as the sunspot number's peak values are concerned, the projection error is given in Table 6. The approach may be compared to the

Table 5. $z(t)$ 11-year projection

year	$z(t)$		year	$z(t)$
	actual	projected		
1990	142.65	127.6	1999	53.4
1991	145.2	134.6	2000	111.9
1992	94.3	91.0	2001	107.9
1993	54.5	53.9	2002	116.8
1994	29.9	29.1	2003	75.1
1995	17.5	12.4	2004	40.0
1996		8.0	2005	14.1
1997		0	2006	0
1998		0		

Table 6. Peak projection error

cycle no.	error %	cycle no.	error %
9	+4.7	16	+2.1
10	+1.7	17	+2.1
11	+1.3	18	+1.5
12	-3.6	19	+8.6
13	-2.8	20	+2.1
14	-3.9	21	+3.7
15	+3.1	22	+7.2

present lean state of the sunspot prediction art as conducted at

the 1996 NASA “consensus panel” convened under Jo Ann Joselyn of the National Oceanic and Atmospheric Administration for the purpose of predicting the maximum of the next sunspot cycle: the panel considered 28 prediction methods of all kinds – including numerology-based ones – and settled on a maximum for the upcoming sunspot cycle no. 23 equal to that of the ongoing 22nd one. Considering that sunspot peaks are accompanied by the emission of highly energetic particles and magnetic storms that can blast the delicate electronics of military, scientific, and commercial satellites, and that the ultraviolet radiation from an active sun heats up Earth’s upper atmosphere causing it to expand and exert extra drag on satellites, a real barn-burner of a maximum could create the need of new shuttle missions to reboost the orbiting Hubble Space Telescope and the planned space station” (Glanz, 1997).

6 Conclusions

The analysis detailed in the foregoing sections brought out and modeled three main features of long-term solar activity based on the sunspot number as the activity’s surrogate: a rising trend, long-term surging, and a patterned transient. Although the idea of using the sunspot number to gauge solar activity is not new (Brown, 1974), the casting of this activity in the role of a long-term dynamic surging agent is. Noteworthy is the fact that the model points to the present decade being the time of peaking of the latest surge S_{III} of long-term solar activity to date, after which a span of roughly five decades of gradual subsiding is to follow. It is also significant that the macrodynamics of a phenomenon as complex as the energy exchanges taking place in the sun’s interior can be portrayed in terms of a representation as simple and compact as the mediating factor having the characteristics of a trigger mechanism. This concept, coupled with the system-response tool of the Kolmogorov formulation, make for an effective analytical tool for the study of at least the phenomenology of sun-related long-term terrestrial phenomena, as Part II of this study will demonstrate. It should be pointed out that the model’s synthesis made use of long-term periodicities present in particular geophysical topics where solar energy certainly has a role. Beyond this, the mediating factor may, as an operational concept, have a bearing upon the search for the modus operandi of the complex dynamics of the interior of the sun itself (Cox *et al.*, 1993). Finally, particular attention of the Kolmogorov algorithm by the applied scientist in any field is warranted; this because of its premise that the dynamics of the studied phenomenon is contained in the historical data itself, even if no theory exists about such dynamics at the present time.

Acknowledgements. The Editor-in-chief thanks two referees for their help in evaluating this paper.

References

- Brown, G. N., A new solar-terrestrial relationship, *Nature*, **251**, 592–594, 1974.
- Clark, J. S., The effect of climate change in fire regimes in northwestern Minnesota, *Nature*, **334**, 233–235, 1988.
- Cohen, T. J., and P. R. Lintz, Long-term periodicities in the Sun, *Nature*, **250**, 398–399, 1974.
- Cox, A. N., W. Livingston, and M. Matthews (Eds.), *Solar interior and atmosphere*, Arizona University Press, Tucson, 1993.
- Dicke, R. K., Is there a chronometer hidden deep in the Sun? *Nature*, **276**, 676–680, 1978.
- Eddy, J. A., The Maunder minimum, *Science*, **192**, 1189–1202, 1976.
- Eddy, J. A., and M. Stuiver, The case of the missing sunspots, *Sci. Am.*, May, 88–92, 1977.
- Friis-Christensen, E., and K. Lassen, Length of the solar cycle: an indicator of solar activity closely associated with climate, *Science*, **254**, 698–700, 1991.
- Fyodorov, M. V., V. V. Klimenko, and V. V. Dovgalyuk, Sunspot minima dates: a secular forecast, *Solar Phys.*, **165**, 193–99, 1996.
- Glanz, J., Bold prediction downplays the Sun’s next peak, *Science*, **275**, 927, 1997.
- Gribbin, J. R., and S. Plegemann, *The Jupiter effect*, Walker, New York, 1974.
- Hoyt, D. V., K. H. Schatten, and E. Nesmes-Ribes, The one hundredth year of Rudolf Wolf’s death: do we have the correct reconstruction of solar activity? *Geophys. Res. Lett.*, **21**, 2067–2070, 1994.
- Hughes, D. W., The inconsistent Sun, *Nature*, **266**, 405–406, 1977.
- James, H. M., N. R. Nichols, and R. S. Phillips (Eds.), *Theory of servomechanisms*, McGraw Hill, New York, 1947.
- Jose, P. D., Sun’s motion and sunspots, *Astron. J.*, **70**, 193–200, 1965.
- Keyfitz, N., Science fragmented, *Options Int. Inst. Adv. Syst. Anal.*, 5–12, 1992.
- Keyfitz, N., On the sectarianism of science, *Options Int. Inst. Adv. Syst. Anal.*, 1–2, 1994.
- Kiepenheuer, K., *The Sun*, University of Michigan Press, Ann Arbor, 1959.
- Kolmogorov, A. N., Interpolation und Extrapolation von stationären zufälligen Folgen, *Bull. Acad. Sci. U.S.S.R.*, Ser. Math. **5**, 1941.
- Lamb, H. H. Volcanic dust in the atmosphere: with a chronology and assessment of its meteorological significance, *Phil. Trans. R. Soc. Lond.*, **266**, 425–527, 1970.
- Lotka, A. J., *Elements of mathematical biology*, Dover, New York, 1956.
- May, R. M., Simple mathematical models with very complicated dynamics, *Nature*, **261**, 474–484, 1976.
- McDonald, M., Nonlinear dynamics, *Nature*, **283**, 431–432, 1980.
- McKinnon, J. A., Sunspot numbers: 1610–1985, *Rep. UAG-95*, World Data Center A, National Academy of Sciences, Washington DC, 1987.
- McNish, A. G., and J. V. Lincoln, Table of observed and predicted relative sunspot numbers and solar flux, *Solar Geophys. Data Suppl.*, 9–14, 1987.
- Modis, T., and A. Bebecker, Chaos-like states can be expected before and after a logistic growth. *Technol. Forecast. Social Change*, **41**, 111–120, 1992.
- Mullin, T., and T. J. Price, Experimental observation of chaos arising from the interaction of steady and time-dependent flows, *Nature*, **340**, 294–296, 1989.
- Plantico, M. S., T. R. Karl, G. Kukla, and J. Gavin, Is recent climate change across the United States related to rising levels of anthropogenic greenhouse gasses? *J. Geophys. Res.*, **95**, 16617–16637, 1990.
- Samuel, Sir H. L., *Essay in physics*, Harcourt, New York, 1952.

- Schneider, S. M., and C. Mass** Volcanic dust, sunspots, and temperature trends, *Science*, **190**, 741–746, 1975.
- Stuiver, M., and P. D. Quay**, Changes in atmospheric carbon-14 attributed to a variable sun, *Science*, **207**, 11–19, 1980.
- Williams, G. E.**, Sunspot periods in the late Precambrian glacial climate and solar-planetary relations, *Nature*, **291**, 624–628, 1981.
- Williams, G. E., and C. P. Sonnett**, Solar signature in sedimentary cycles from the Elatina formation, Australia, *Nature*, **318**, 523–527, 1985.
- Wilson, R. M.**, On the use of “first spotless day” as a predictor for sunspot minimum, *Solar Phys.* **158**, 197–204, 1985.
- Withbroe, G. L.**, Solar activity cycle: history and prediction, *J. Spacecraft*, **26**, 394–402, 1989.
- Wood, K. D.**, Sunspots and planets, *Nature*, **240**, 91–93, 1972.