Solar resonant diffusion waves as a driver of terrestrial climate change

Robert Ehrlich*

George Mason University, Fairfax, VA 22030, USA

Received 11 September 2006; received in revised form 4 January 2007; accepted 9 January 2007

Abstract

A theory is described based on resonant thermal diffusion waves in the sun that explains many details of the paleotemperature record for the last 5.3 million years. These include the observed periodicities, the relative strengths of each observed cycle, and the sudden emergence in time for the 100 thousand year cycle. Other prior work suggesting a link between terrestrial paleoclimate and solar luminosity variations has not provided any specific mechanism. The particular mechanism described here has been demonstrated empirically, although not previously invoked in the solar context. The theory, while not without its own unresolved issues, also lacks most of the problems associated with Milankovitch cycle theory.

Keywords: Diffusion waves; Modelling; Paleoclimate

1. Paleotemperature data

The possibility that fluctuations in solar luminosity may be responsible for changes in global temperatures has not been overlooked by researchers (Lean, 1997), although most explanations for periodicities in paleotemperatures are believed to involve factors unrelated to solar luminosity (Zachos et al., 2001). Nevertheless, some researchers have suggested that periodic solar variability has been the cause of global temperature cycles, with periods ranging from the 11 year sunspot period to cycles as long as 2000 years (Hu et al., 2003; Van Geel, 1999). Earlier reports, however, suggest no specific mechanism for long term solar periodicity.

Paleotemperatures are inferred from the $^{18}\text{O}/^{16}\text{O}$ ratio found in benthic sediments and in ice cores versus depth. In particular, it is believed that a 5°C decrease in global temperature corresponds to a 0.1% increase in the $^{18}\text{O}/^{16}\text{O}$ ratio (i.e., delta $^{18}\text{O}$) in benthic sediments (Lisiecki and Raymo, 2005). Fig. 1 shows a compilation by Lisiecki and Raymo (2005) who stacked 57 sets of data on benthic sediments that cover 5.3 million years (My) before present. Figs. 2 and 3 show the Fourier amplitude spectra.

2. Diffusion waves

Diffusion waves are an established phenomenon which has been studied theoretically and experimentally (Mandelis, 2000), and even for heat transport in magnetized plasmas (Reynolds et al.,...
They can occur for any process, such as heat flow when a time-dependent source is present, in which case, the waves represent propagating temperature fluctuations, \( c(r, t) \), and they must satisfy the diffusion equation. Assuming a sinusoidal time dependence, and separating \( c(r, t) = f(r)e^{i\omega t} \) into space and time-dependent parts, we find that \( f(r) \) satisfies

\[
\nabla^2 \phi - \kappa^2 \phi = Q_0(r), \tag{1}
\]

where \( Q_0(r) \) is the space-dependent part of a source term, and where the complex wave number \( \kappa \) satisfies

\[
\kappa = \sqrt{i\omega} \frac{1}{D_t} = (1 + i) \sqrt{\frac{\omega}{2D_t}}. \tag{2}
\]

Note that according to Eq. (2), the imaginary (attenuation) part of the complex wave number \( \kappa \) and the real (oscillatory) part are equal, implying equality of penetration depth \( L = 2\pi \text{Re}(\kappa) \), and wavelength \( \lambda = 2\pi / \text{Re}(\kappa) \), so higher frequency waves are more severely damped. For a uniform thermal diffusivity \( D_t \) and a spherically symmetric source Eq. (1) yields the solution

\[
\phi(r) = \frac{\sin(Re(\kappa)r)}{r}. \tag{3}
\]
For the more general case of an isotropic medium in which $k(r)$ is dependent only on radial distance to the origin the same solution applies, but $k$ must now be replaced by its average value between the source radius $r_0$ and the $r$ at the wave location, i.e.,

$$\bar{k}(r - r_0) = \int_{r_0}^{r} k(r) \, dr. \quad (4)$$

3. The solar interior

Solar physicists use a Standard Solar Model (SSM) to calculate many properties of the sun’s interior. According to the SSM, the three interior regions of interest here are: (1) the core (of radius about 25% of the solar radius $r_\odot$) where fusion occurs, (2) the surrounding radiative zone (extends to about 0.7$r_\odot$), and (3) the surrounding convection zone, where heat makes its way to the surface by the much more rapid process of convective flow. The fairly sharp boundary layer between the radiative and convection zones is known as the “tachocline.”

Normally, one associates diffusivity with heat flow by conduction, whereas radiation is taken as a line-of-sight transmission process. However, inside the sun where the mean free path for photons between emission and subsequent reabsorption is very short (typically several millimeters), the diffusion description is applicable. Obviously, radiation is the dominant heat transfer process inside the core and radiative zones, owing to the high temperatures there, and the $T^4$ dependence on radiative emission and absorption rates. The diffusivity parameter depends on a variety of other quantities that govern the likelihood of photons being radiated and reabsorbed. In particular, it can be shown that the thermal diffusivity at any radial distance as (Miesch, 2005): $D_t = 16\sigma_{sb} T^3/(3\chi \rho^2 C_P)$, where $\sigma_{sb}$ is the Stefan–Boltzmann constant, $T$ is the absolute temperature, $\chi$ is the opacity, $\rho$ is the density, and $C_P$ is the specific heat. Each of the parameters in the previous equation can be found as a function of radial distance using data from the GONG collaboration (Christensen-Dalsgaard, 1996).

Within the SSM description, the solar core is considered to be in a state of quasistatic equilibrium changing only on a timescale of 30 million years. However, Grandpierre and Agoston (2005) have noted, instabilities can be generated in a rotating plasma in the presence of a magnetic field, and these instabilities can give rise to thermal fluctuations, and heat waves that represent deviations from SSMs. They further show that constraints on the size of the core magnetic field (Friedland and Gruzinov, 2004), (Gough and MacIntyre, 1998) are such that fluctuations can indeed arise, as a natural consequence of the nonlinearity of the MHD equations.

These thermal fluctuations in the form of hot “bubbles” tend to expand, rise, and cool on...
timescales that Grandpierre and Agoston (2005) have estimated, which are a function of their spatial extent and the magnitude of their temperature excess. Were the bubbles to remain small and isolated, the diffusion waves emanating from these fluctuations would tend to cancel each other out. However, Grandpierre and Agoston (2005) have shown that the bubbles tend to merge and grow in size, and the decay time of the waves can grow even longer than $10^7$ years, as the perturbations smooth out to include ever larger regions—in effect making the solar core a region of small amplitude thermal oscillations, instead of the quiescent state normally assumed.

4. Description of the theory

We postulate small random variations in the temperature of the outer solar core (between $0.21 r_{\odot}$ and $0.25 r_{\odot}$) (Ferro et al., 2005), which are sufficiently smoothed out, so as to be a function only of radial distance. The specific noise spectrum is unimportant, except that the noise frequencies should have slowly changing phases over times on a scale of $10^5$ y.

Each of the noise frequencies may be thought of as a source of radially symmetric diffusion waves. When a wave reaches the tachocline, it is reflected back to the core as an ingoing wave. However, it should be noted that reflections in the case of diffusion waves are not of the usual type, where well-defined wavefronts are involved, but only resemble conventional wave reflections in the high frequency or long path-length limit (Mandelis et al., 2001). Otherwise, a considerable amount of “phase smearing” prevents us from describing the reflection of discrete wavefronts. At a boundary, such as the tachocline, the depletion or accumulation of photons (and the associated radially symmetric periodic temperature fluctuation), creates an inward travelling diffusion wave. This inward wave interferes with the next outgoing wave produced at that frequency, and a new outward wave is generated from a superposition of the two waves. Again, however, the interference process is not one of simply superimposing well-defined wavefronts, which do not exist here. Instead, if the inward and outward waves both correspond to a maximum positive or negative fluctuation in photon density (i.e., have matched phases), the resultant temperature fluctuation would be greatest, and hence resonance would be achieved. To achieve resonance the ingoing wave must arrive back at the core in phase with the next source cycle at this frequency, or more generally after $n$ cycles. Thus, the condition for resonance for the $n$th mode is

$$\text{Re}(\varepsilon)2r = 2r \sqrt{\frac{\omega}{2D_t}} = 2\pi n,$$  \hspace{1cm} (5)

where $n = 1, 2, 3, \ldots$, and $2r$ is the round trip distance travelled by the wave from emission to absorption back to the core. From Eq. (5) we see that the period $T_n$ of the $n$th resonant mode $T_n = 2\pi/\omega_n \propto 1/n^2$, so that

$$T_n = T_1/n^2,$$  \hspace{1cm} (6)

with the period $T_1$ of the lowest frequency resonant mode being twice the diffusion time $t_d$ for photons to undergo a random walk through the radiation zone before they reach the tachocline. Fortunately, $t_d$ can be calculated from the sun’s internal density and opacity (both functions of radial distance $r$) using data from the SSM (Christensen-Dalsgaard, 2005). One of the most widely cited estimates of $t_d$ from Mitalas and Sills (1992) is $1.7 \times 10^6$ y = 170 ky. We have repeated their calculation, and obtained 190 ky. In what follows, we shall use a value that brackets these two: $t_d = 180 \pm 10$ ky.

Using Eq. (6) we find the predicted periods $T_n$ for the $n$th resonant mode:

$$T_n = \frac{360 \pm 20}{n^2} \text{ ky}.$$  \hspace{1cm} (7)

Some of the periods (for $n = 2, 3$ and 4) given by Eq. (7) correspond to reported periods in the paleotemperature data—see markers in Fig. 2. But before we can have confidence that resonant solar diffusion waves may offer an explanation of periodic variations in paleotemperatures, we must address several key issues, including whether it is possible for the waves to attain a large degree of resonant amplification.

4.1. Damping of diffusion waves

Given that a diffusion wave for the $n$th resonant mode has a phase of $2\pi n$ on returning to the source, it will have been heavily damped, and will have an intrinsic damping:

$$d_n = e^{-2\pi n}.$$  \hspace{1cm} (8)

The large damping of diffusion waves might seem to make the possibility of their playing a role in any...
resonant amplification process in the sun remote. For example, given a damping $d_n$ for a wave after one round trip, the next outgoing wave has an amplitude $1 + d_n$, and hence the number of cycles required for the amplitude of the $n$th mode to double is given as $\ln 2/d_n$, yielding doubling times in kyr for mode $n$: $t_{2n} = 2n^2 T_{\text{eff}} e^{2n} \ln 2$. For $n > 1$ we find doubling times that are longer than the age of the sun—even if diffusion waves are reflected 100% from the tachocline. Yet the doubling times are lengthened even more, because of damping due to “blurry boundaries,” since fractions of an outgoing wave are reflected back from different parts of the tachocline, and return to different parts of the core.

Based on the relevant layer thicknesses, we find a phase smearing of the $n$th mode: $\Delta \phi \approx \pm 2\pi n/10$. This additional damping, which increases with $n$, makes it impossible for modes $n > 4$ to survive.

### 4.2. Gain due to fusion amplification

There is a compensating amplification effect that more than offsets the enormous damping and allows resonant diffusion waves to emerge with large amplitude in a relatively few cycles. According to the SSM, the rate of nuclear fusion in the core of the sun (and hence solar luminosity $L$) is proportional to $T^{4.2}$ for small deviations from the ambient core temperature (Bahcall et al., 2001). To see the effect of this large exponent in amplifying diffusion waves, consider an outgoing wave for mode $n$ initially created due to a positive core temperature fluctuation $\phi_0(r_0)$ at radius $r_0 = 0.21r_\odot$. In advancing a small distance $dr$ outward, the wave amplitude is reduced by $d\phi = kdr$. However, if the wave is absorbed by the core, it elevates the core temperature at that $r$ by the same amount, and hence raises the luminosity there by a 4.2-fold greater fraction, due to the $L \propto T^{4.2}$ relationship. Thus, there is a net gain in wave amplitude over this distance $dr$ of $3.2 \pi dr$. The amplification factor multiplies the size of the temperature fluctuation, and it occurs virtually instantaneously on a diffusion time scale, so that the amplification looks like a coherent radiation wave superimposed on a much slower moving thermal wave of the same frequency. As the wave proceeds outward through the outer core, it continues to produce an ever increasing amount of thermal radiation, owing to the continual increase in luminosity. Upon reaching the edge of the core ($0.25r_\odot$), the $n$th resonant mode (and the core temperature) has been amplified by a gain

$$g_n = e^{3.2 \pi n r},$$

where $r$ can be found using Eqs. (2), (4), (5), with $r_0 = 0.21r_\odot$ and $0.25r_\odot$. The same mechanism operates for the negative portion of the wave cycle, and it amplifies reductions in core temperature. This amplification mechanism is similar to the exponential increase in intensity in the case of lasing, except that here (1) fusion makes optical pumping unnecessary, and (2) the spherical symmetry assures that there is no loss of waves “out the sides,” as occurs in a laser. Evaluating Eq. (9) numerically using SSM data yields $g_n = e^{0.6n}$.

Such a large gain (more than offsetting the $e^{-2\pi n}$ damping) cannot be taken as a realistic indicator of how large the wave can grow before it leaves the core, but it does suggest that a wave might be able to resonantly attain its maximum value after only a few cycles. It should be noted that a very similar resonance phenomenon to that described here—but without any fusion gain—has been seen empirically in thermal-wave cavity environments (Shen and Mandelis, 1995).

Obviously, the theory provided here, which ignores non-thermal magnetic interactions, and assumes an initial temperature fluctuation that is only a function of radial distance is highly simplified. Far more numerous non-spherically symmetric temperature fluctuations would not result in a similar amplification process, and hence they will become relatively less intense over time. However, both the non-spherically symmetric fluctuations and the magnetic interactions must be considered for a full understanding of losses and gains, so as to permit a determination of the maximum steady state radial wave amplitude.

### 5. Comparison of between theory and paleoclimate data

#### 5.1. Periods of each mode

As Fig. 2 (markers) and Table 1 show, the Fourier amplitude spectrum for the data shows peaks at the theory’s predicted modes $n = 2, 3$ and 4. The overall data with a fifth order background subtraction to eliminate low frequency leakage shows no indication of the $n = 1$ mode in Fig. 2. With no background subtraction there is a non-statistically significant suggestion of a peak corresponding to this mode in Fig. 3(c) and especially (d), which together correspond to the interval

Please cite this article as: Ehrlich, R., Solar resonant diffusion waves as a driver of terrestrial climate change, Journal of Atmospheric and Solar-Terrestrial Physics (2007), doi:10.1016/j.jastp.2007.01.005
Table 1

<table>
<thead>
<tr>
<th>Mode n</th>
<th>$T_n$ Pred</th>
<th>$T_n$ Obs</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>360 ± 20</td>
<td>—</td>
</tr>
<tr>
<td>2</td>
<td>90 ± 5</td>
<td>95 ± 10</td>
</tr>
<tr>
<td>3</td>
<td>40 ± 2</td>
<td>41.0 ± 0.4</td>
</tr>
<tr>
<td>4</td>
<td>22.5 ± 1.5</td>
<td>(23, 7, 22.4)</td>
</tr>
</tbody>
</table>

Uncertainties on the observed values are based on the half widths of the peaks in Fig. 2.

2048–4095 y before present. Were it present, a $n = 1$ peak (located at $f = 0.28$ cycles per 100 ky) could not easily be distinguished from a steeply rising low frequency background.

No periodicities predicted by the model are clearly absent from the data. It is true, however, that one small peak seen in the overall spectrum (but not in the quartile spectra) with a period of 19 ky (dotted arrow in Fig. 2) is not predicted by the theory, and it cannot be the $n = 5$ mode which should be at 14.8 ky were its predicted amplitude is not zero. Obviously, we cannot exclude the possibility that other mechanisms, including Malinkovitch cycles, play some role in paleotemperature periodicity.

5.2. Signal emergence times

By the signal "emergence time" we refer to the amount of time required for the $n$th mode amplitude to rise from the modest level of the background noise to a large enough multiple for it to be clearly seen in the paleotemperature record. Owing to the large gain, this amplification could occur in a few cycles according to the theory. We can determine the emergence time from the data only for modes that turn on after being off for a while. Only one of the modes (i.e., $n = 2$ with a 90–100 ky period) clearly shows this behavior, since as seen in Fig. 3, it is present only in the first and third time quartiles (solid arrows). It is not easy to tell from Fig. 1 exactly where in the vicinity of 1000 ky before the present this mode begins to appear, and just how sudden is that appearance.

The shape of the $n = 2$ cycle, and the abruptness of its emergence can be gleaned by looking at the data for the most recent 1000 ky, after making a subtraction of the 41 ky cycle—see Fig. 4. The figure also shows vertical error bars on the original data, and a 100 ky cycle drawn to guide the eye. When the data is displayed in this manner the 100 ky cycle appears much more consistent with a sinusoidal shape than the original data for the last 1000 ky.

It can easily be seen in Fig. 4 that the $n = 2$ mode appears fairly suddenly (between 800 and 900 ky before the present), following what appears to be a sudden phase reversal or discontinuity. If this discontinuity is real, the theory has no obvious explanation for it. However, a short emergence time is consistent with the theory, so this behavior offers some support.

5.3. Relative amplitude (and shape) of each mode

We have addressed the issue of the lack of a clear signal for the $n = 1$ mode, and the non-appearance of modes with $n > 4$. Can the theory account for the relative amplitude (and shape) of the three other modes ($n = 2, 3, 4$) that are clearly seen in the data? It can with the aid of two further assumptions: (1) all modes have the same intrinsic strength (as a multiple of the background) when they are active, and (2) only the $n = 3$ mode is active most of the time.

$n = 2$ mode (90 ky period): A look at Figs. 3(a) and (c), i.e., times when both the $n = 2$ and 3 modes were active, shows that the two peaks have comparable strengths. (The position of the $n = 2$ mode is flagged by solid and dotted arrows for those quartiles that it appears to be “on” and “off”).

Thus, the decreased size of this mode in the overall data is apparently less, a matter of the differences in the intrinsic strength of each mode than what
fraction of the time this mode is active. The broadening of this peak may also be attributed to its being off perhaps half the time.

\[ n = 3 \text{ mode (40 ky period):} \] This mode is the dominant (and narrowest) one in the overall record because it is active at all times—see Figs. 3(a)–(d).

\[ n = 4 \text{ mode (22.5 ky period):} \] This mode is not clearly seen in the separate quartile spectra (Figs. 3(a)–(d)). However, the fact that it appears to be split in the overall spectrum (into two frequency peaks separated by about 7%) has a possible explanation that relates to a single frequency turning on and off in a periodic manner. For example, if we perform a Fourier Transform on a sinusoidal signal that turns off and on every 14 cycles, we would indeed get two separate peaks in the frequency domain separated in frequency by 7%. This phenomenon is not due to a defect of the Fourier Transform, but rather that two nearby frequencies will produce beats which mimic a single frequency turning on and off periodically. It may simply be a coincidence, but this number of cycles on and off for the \( n = 4 \) mode is close to what is seen for the \( n = 2 \) mode, which appears to be alternately on and off for approximate intervals of 1028 ky (≈ 11 cycles).

Finally, how to explain the smallness of the \( n = 4 \) peak compared to \( n = 3 \)? Its splitting into two separate peaks is partly responsible, as is the fact that this mode is active only a fraction of the time, unlike the \( n = 3 \) mode. A third reason for the lower \( n = 4 \) amplitude is that the damping of this mode due to blurry boundaries (twice as serious for \( n = 4 \) than 3) operates mainly as the amplification process occurring while the wave travels through the core, which has the effect of preventing the \( n = 4 \) mode from reaching a large amplitude.

5.4. Can the oscillations become large enough?

The dominant (\( n = 3 \)) mode seen in the Fourier spectrum (Fig. 2) is seen to have a signal to background ratio of about 3.4. This implies that the background noise at that frequency had to be resonantly amplified by this same factor. Given that the complex interactions with the magnetic field that might determine an upper limit to the size of the oscillations remains to be worked out, it is unclear if a 3.4-fold growth (at this particular frequency) is achievable. This unresolved issue does represent a current shortcoming of the theory.

6. Problems with Milankovitch theory and conclusion

In Milankovitch theory past glaciations are assumed to arise from small quasiperiodic changes in the Earth’s orbital parameters that give rise to corresponding changes in solar insolation, particularly in the polar regions. A brief discussion of five problems with this theory are listed below, and a more detailed description of some of them can be found elsewhere (Karner and Muller, 2000; Bolfshakov, 2003).

(a) Weak forcing problem: The basic problem with the theory is that the observed climate variations are much more intense than the insolation changes can explain without postulating some very strong positive feedback mechanism.

(b) 100 ky problem: The preceding basic problem can be illustrated for the case of one particular parameter—the orbital eccentricity. The dominant climate cycle observed during the last million years has a roughly 100 ky period, which in Milankovitch theory is linked to a 100 ky cycle in the eccentricity. However, the effect of this eccentricity variation should be the weakest of all the climate-altering changes, in view of the small change in solar insolation it would cause. For example, consider the Earth’s orbital eccentricity, \( e \), which has been shown to have several periods including one of 100 ky during which \( e \) varies in the approximate range: \( e = 0.03 \pm 0.02 \) (Quinn et al., 1991). The resultant solar irradiance variation found by integrating over one orbit for each of the two extreme \( e \) values is about \( \pm 0.05\% \), or \( \pm 0.17 \) W/m\(^2\) difference at the top of the Earth’s atmosphere. Given that climate models show that a 1% change in solar irradiance would lead to a 1.8°C average global temperature change, then the change resulting from a \( \pm 0.05\% \) irradiance change would be a miniscule 0.1°C hardly enough to induce a major climate event—even with significant positive feedback.

(c) 400 ky problem: The variations in the Earth’s orbital eccentricity show a 400 ky cycle in addition to the 100 ky cycle, with the two cycles being of comparable strength. Yet, the record of Earth’s climate variations only shows clear evidence for the latter.

(d) Causality problem: Based on a numerical integration of Earth’s orbit, a warming climate predates by about 10,000 years the change in insolation than supposedly had been its cause.

(e) Transition problem: No explanation is offered for the abrupt switch in climate periodicity from 41...
to 100 ky that is found to have occurred about a million years ago. Of these five problems with Milankovitch theory, the current theory clearly shares only (c).

In conclusion, We have here suggested a specific mechanism involving diffusion waves in the sun whose amplitude should grow very rapidly due to an amplification provided by the link between solar core temperature and luminosity. Moreover, the phenomenon of resonant amplification of thermal diffusion waves has been empirically demonstrated, albeit not in the solar context (Shen and Mandelis, 1995). A number of features of the theory still remain to be resolved, but the theory does explain many features of the paleotemperature record, and it appears to be free of most defects of the Milankovitch theory. The theory further implicitly suggests the existence of a new category of variable stars having extremely long periods—i.e., \( \approx 10^4 \) times longer than stellar periods currently considered to be “very long.” For some stars with \( M < M_\odot \), their thinner radiation zones might make the predicted periods observable.

7. Uncited references

Remhof et al., 2003.

References


According to the SSM the CNO cycle contributes 1.5% of the sun’s luminosity, with the remainder due to the pp chain. At the ambient solar core temperature, these reactions have a luminosity dependence on core temperature that varies as of the form \( T^4 \), where \( N \) is approximately 4 and 15, respectively. To find the overall dependence on temperature, we may compute numerically the derivative \( d\phi/dT \) after weighting the two reactions by their respective abundances, giving \( N = 4.2 \).


Ferro, F., Lavagno, A., Quarati, P., 2005. Metastable and stable equilibrium states of stellar electron-nuclear plasmas. Physics Letters A336, 370–377 Ferro and Lavagno have shown that it is not the entire core of the sun that is in a metastable state, but only its outer region, i.e., \( r > 0.21r_\odot \).


Please cite this article as: Ehrlich, R., Solar resonant diffusion waves as a driver of terrestrial climate change, Journal of Atmospheric and Solar-Terrestrial Physics (2007), doi:10.1016/j.jastp.2007.01.005