Celestial Climate Driver: A Perspective from Four Billion Years of the Carbon Cycle

Ján Veizer
Ottawa-Carleton Geoscience Centre, University of Ottawa, Ottawa, K1N 6N5 Canada & Institut für Geologie, Mineralogie und Geophysik, Ruhr-Universität Bochum, Bochum, Germany: veizer@science.uottawa.ca.

SUMMARY
The standard explanation for vagaries of our climate, championed by the IPCC (Intergovernmental Panel on Climate Change), is that greenhouse gases, particularly carbon dioxide, are its principal driver. Recently, an alternative model that the sun is the principal driver was revived by a host of empirical observations. Neither atmospheric carbon dioxide nor solar variability can alone explain the magnitude of the observed temperature increase over the last century of about 0.6°C. Therefore, an amplifier is required. In the general climate models (GCM), the bulk of the calculated temperature increase over the last century of about 0.6°C. Therefore, an amplifier is required. In the general climate models (GCM), the bulk of the calculated temperature increase is attributed to “positive water vapour feedback”. In the sun-driven alternative, it may be the cosmic ray flux (CRF), energetic particles that hit the atmosphere, potentially generating cloud condensation nuclei (CCN). Clouds then cool, act as a mirror and reflect the solar energy back into space. The intensity of CRF reaching the earth depends on the intensity of the solar (and terrestrial) magnetic field that acts as a shield against cosmic rays, and it is this shield that is, in turn, modulated by solar activity.

Cosmic rays, in addition to CCN, also generate the so-called cosmogenic nuclides, such as beryllium-10, carbon-14 and chlorine-36. These can serve as indirect proxies for solar activity and can be measured e.g., in ancient sediments, trees, and shells. Other proxies, such as oxygen and hydrogen isotopes can reflect past temperatures, carbon isotopes levels of carbon dioxide, boron isotopes the acidity of ancient oceans, etc. Comparison of temperature records from geological and instrumental archives with the trends for these proxies may enable us to decide which one of the two alternatives was, and potentially is, primarily responsible for climate variability. This, in turn, should enable us to devise appropriate countermeasures for amelioration of human impact on air quality and climate.

SOMMAIRE
Généralement, les raisons données pour expliquer les caprices de notre climat, les mêmes que celles avancées par le CICC (Comité intergouvernemental sur le changement climatique), veulent que ces soient les gaz à effet de serre, particulièrement le dioxyde de carbone, qui en soient le moteur principal. Récemment, une série d’observations empiriques ont ravivé l’intérêt pour un autre modèle voulant que ce soit le soleil qui en soit le moteur principal. Mais seuls, ni le dioxyde de carbone ni les variations d’activité solaire ne permet d’expliquer la hausse de température observée au cours du siècle dernier, soit environ 0,6 °C. D’où la nécessité d’un facteur d’amplification. Dans les modèles climatiques généraux (GCM), le gros de l’accroissement calculé de température est dû à « la rétroaction positive de la vapeur d’eau ».

Dans le modèle à moteur solaire, ce pourrait être le flux de rayonnement cosmique (FRC), ce pourrait être l’effet des particules énergiques qui en frappant l’atmosphère entraînent une génération possible de nucléides de condensation des nuages (NCN). Alors, les nuages se refroidissent et, comme un miroir, réfléchissent l’énergie solaire dans l’espace. L’intensité du FRC atteignant le sol dépend de l’intensité des champs magnétiques du soleil et de la Terre, lesquels agissent comme un bouclier à l’endroit des rayons cosmiques, le pouvoir de ce bouclier étant à son tour modulé par l’activité solaire. En plus d’entraîner la formation de NCN, les rayons cosmiques, génèrent aussi ce qu’on appelle des nucléides cosmogéniques, comme le beryllium-10, le carbone-14 et le chlore-36. Ces nucléides peuvent servir d’indicateurs indirects de l’activité solaire puisqu’on peut en mesurer la teneur dans des sédiments anciens, des arbres, et des coquilles, par exemple. D’autres indicateurs indirects comme les isotopes d’oxygène et d’hydrogène peuvent refléter les températures de jadis, les isotopes de carbone peuvent refléter les niveaux de dioxyde de carbone, les isotopes de bore peuvent refléter l’acidité des anciens océans, etc. La comparaison entre des registres de mesures de température directes et d’archives géologiques, avec les courbes de tendance de tels indicateurs indirects peut nous permettre de décider laquelle de deux options était et continue possible-ment d’être la cause principale des varia-
tions climatiques. On pourrait alors décider de contre-mesures appropriées permettant d’atténuer l’impact des activités humaines sur la qualité de l’aire et sur le climat.

**INTRODUCTION**

Carbon dioxide, generally believed to be the most important greenhouse gas and climate modifier, is today the focus of a heated political and scientific debate that has polarized scientists, policy makers, and the public. One side maintains that CO₂ is the principal driver of climate, with the Intergovernmental Panel on Climate Change (IPCC, 2001) projecting a global mean temperature rise from 1.5 to 5.8° C by the year 2100. The other side (e.g., Douglass et al., 2004) claims that the role of anthropogenic CO₂ on climate has not been proven, and that there is therefore no need for emissions quotas such as those mandated by the Kyoto Protocol.

As is usually the case with contentious matters, the reality likely lies somewhere in between. So why is this issue so polarizing? First, past, natural, variations in the carbon cycle and climate modifier, is today the focus of a heated political and scientific debate that has polarized scientists, policy makers, and the public. One side maintains that CO₂ is the principal driver of climate, with the Intergovernmental Panel on Climate Change (IPCC, 2001) projecting a global mean temperature rise from 1.5 to 5.8° C by the year 2100. The other side (e.g., Douglass et al., 2004) claims that the role of anthropogenic CO₂ on climate has not been proven, and that there is therefore no need for emissions quotas such as those mandated by the Kyoto Protocol.

**CELESTIAL CLIMATE DRIVER**

The solar/Cosmic Ray Flux (CRF)/climate hypothesis, although discussed by the IPCC (Ramaswamy et al., 2001), was not considered to be a likely candidate for a principal climate driver. This was partly because of the lack of a robust physical formulation for cloud condensation phenomena and partly because it was argued that the observed changes in the Total Solar Irradiance (TSI) flux have been insufficient to account for the observed ~0.6° C centennial temperature increase. Therefore, an amplifier is required to account for the discrepancy. However, similar problems have arisen also in the greenhouse hypothesis, where the amplifier is implicit (the centennial temperature rise in these models is caused by the “positive water vapour feedback”, not to the CO₂ itself) and where clouds, a potential net negative feedback and the largest source of uncertainty in the models, are only “parameterized”. Yet, the solar energy reflected by the clouds, or the energy of evaporation/condensation, are both about 78 Watts per square metre (Wm⁻²) worldwide. For comparison, the energy input ascribed to “post-industrial” anthropogenic CO₂ input is ~ 1.5 Wm⁻² and that of incoming solar radiation ~ 342 Wm⁻² (IPCC, 2001). A change in cloud cover of a few percent can therefore have a large impact on the planetary energy balance.

A growing body of empirical evidence, such as correlations between climate records and solar and cosmic ray activity, or their proxy indicators (e.g., ¹⁰⁷Be, ¹⁴C, ³⁶Cl), geomagnetic field intensity, sunspot numbers), increasingly suggests that extraterrestrial phenomena may be responsible for at least some climatic variability (Bond et al., 2001; Kromer et al., 2001; Neff et al., 2001; Sharma, 2002; Carslaw et al., 2002; Hu et al., 2003; Usoskin et al., 2003; Blaauw et al., 2004; Solanki et al., 2004). The correlations of climate with these proxies are mostly better than those, if any, between the coeval climate and CO₂.

Moreover, inferred and direct observational data of TSI flux yield a record that can explain 80% of the variance in the centennial temperature trend (Foukal, 2002). Celestial phenomena may have been the principal driving factor of climate variability and global temperature even in the recent past.

The sun-climate link could be through a number of potential pathways (Rind, 2002; Carslaw et al., 2002), where the solar flux is amplified by (1) stratospheric chemistry (e.g., ozone) because of changes in solar UV spectrum, (2) cloud coverage modulated by the galactic CRF, or (3) a combination of these or other factors. Considering that statistical evaluation of 20ᵗʰ century data shows that solar UV radiation may account for only about 20% of the variance in surface temperature data (Foukal, 2002), alternative (2) is the favoured hypothesis. In this alternative, an increase in TSI results not only in an enhanced thermal energy flux, but also in more intense solar wind that attenuates the CRF reaching the Earth (Tinsley and Deen, 1991; Svensmark and Friis-Christensen, 1997; Marsh and Svensmark, 2000; Solanki, 2002). This, the so-called heliomagnetic modulation effect reflects the fact that the solar magnetic field is proportional to TSI and it is this magnetic field that acts as a shield against cosmic rays. The terrestrial magnetic field acts as a complementary shield, and its impact on CRF is referred to as geomagnetic modulation (Beer et al., 2002). The CRF, in turn, is believed to correlate with the low altitude cloud cover (Fig. 2). The postulated causation sequence is therefore: brighter sun ⇒ enhanced thermal.

![Figure 1. Hierarchy of time scales discussed in this article.](image-url)
flux + solar wind ⇒ muted CRF ⇒ less low-level clouds ⇒ lower albedo ⇒ warmer climate. Diminished solar activity results in an opposite effect. The CRF/cloud-cover/climate link is also physically feasible because the CRF likely governs the atmospheric ionization rate (Carslaw et al., 2002), and because recent theoretical and experimental studies relate the CRF to the formation of charged aerosols (Harrison and Aplin, 2001; Lee et al., 2003), which could serve as cloud condensation nuclei (CCN), as was demonstrated independently by ground based and airborne experiments (Eichkorn et al., 2002).

The CRF reaching the planet has not only an extrinsic variability reflecting its attenuation by solar wind, but also an intrinsic one arising from a variable interstellar environment (Shaviv, 2002a, b). Particularly large CRF variability should arise from passages of the solar system through the Milky Way’s spiral arms that harbour most of the star formation activity. Such passages recur at about 143 ± 10 million years (Ma) intervals and these variations are expected to be about an order of magnitude more effective than the extrinsic ones.

In a nutshell, the intrinsic interstellar intensity of CRF may have controlled the long-term climate variability on multimillion-year time scales. Superimposed on this long-frequency/large-amplitude wavelength are smaller oscillations on millennial to annual time scales, generated by the variable solar activity that modulates either the CRF bombarding the Earth, the planetary atmospheric dynamics, or both. Tentatively, I accept this interpretation as a working hypothesis for the subsequent discussion, but hasten to acknowledge that the CRF/cloud link-age is still a hotly contested issue. Accepting this scenario as a working hypothesis, how does it withstand scrutiny if tested against the hierarchical geomologic record (Fig. 1) of climate and the carbon cycle?

**Figure 2.** Solar irradiance (SI), galactic cosmic ray (CR) flux and low cloud (LC) cover, 1983 – 2001 (adapted from Marsh and Svensmark, 2003a and Marsh et al., 2005). Note the reversed scale for SI. Some authors (Laut, 2003) argue that the apparent post-1995 divergence of clouds from celestial trends disqualifies the correlations. However, the discrepancy may arise from a modified cross-calibration of satellites, following the late 1994 hiatus in polar orbit flights (Marsh and Svensmark, 2003a). A correction for this drift (thick full line LC’) results in a good agreement for all parameters (see also Pallé et al., 2004b and Usoskin et al., 2004).

**Figure 3.** Idealized reconstruction of the oldest, 3.5 billion years (Ga) old, fossils from Western Australia, considered to be blue-green algae (Schopf, 1983). The biogenic origin of these fossils has recently become a matter of controversy (Brasier et al., 2002). Nevertheless, stromatolites and carbon isotope evidence support the great antiquity of life.

**LIFE, WATER, AND THE CARBON CYCLE ON BILLION YEAR TIME SCALES**

To understand the role of atmosphere, water, and life in climate evolution over geologic history, it is essential to study
ancient examples. Yet, we have no unequivocal samples of ancient waters, and the oldest samples of air are in bubbles frozen into Antarctic ice near the time of its formation, reaching back some 420,000 to 800,000 years. The situation is somewhat better with the remnants of life, because mineralized shells go back to about 545 million years, the times known as the Phanerzoic, and morphological evidence of living things, algae and bacteria, and of fossilized stromatolites, have been found in western Australia in rocks as old as 3.5 billion years (Fig. 3). Kerogen, body tissues altered by temperature and pressure, has been found in still older rocks approaching 4 billion years. This is remarkable, because the oldest rocks ever recovered, found near Yellowknife in northwestern Canada, are of about the same age (Bowring et al., 1989).

These observations, however, are only qualitative. If we want to understand the operation of the carbon cycle and its role in the climate system, it is necessary to know not only that there was life, but also how much of it there was. In order to establish this, we have to rely on the derivative, or proxy, signals. In our case, such proxies are isotopes, particularly of carbon and oxygen.

From the measurements of isotope ratios of carbon in modern living things and of carbon dissolved in seawater, the rough proportion of reduced to oxidized carbon is calculated to be about 1:4 (Schidlowski et al., 1975).

Remarkably, when these carbon isotopes are traced back in geologic history, the average carbon isotopic composition of seawater (Fig. 4) and of most of the kerogen (Hayes et al., 1983) was similar to today. Hence, we get about the same 1:4 ratio as far back as 3.5, and possibly 4, billion years ago. Assuming that the stocks of global carbon were conservative, and stated rather boldly, not only did we have life as far back as we had rocks, but there was as much life then as today, albeit in its primitive form. We can conclude, then, that the fundamental features of the carbon cycle were established as early as 4 billion years ago.

What does this mean for the global carbon cycle? The simplest assumption would be that it might not have been that different from today. Yet, such a proposition is difficult to reconcile with the so-called “faint young sun” paradox (Sagan and Mullen, 1972). Based on our understanding of the evolution of stars, the young sun was about 30 percent less luminous than it is today, and became brighter with age. With such low radiative energy from the sun, our planet should have been a frozen ice ball until about 1 billion years (Ga) ago. Yet, we know that running water shaped the surface of the planet as far back as the geologic record goes.

To resolve this paradox, some argue that a massive greenhouse, caused principally by CO₂ (e.g., Kasting, 1993), must have warmed up the young earth. Theoretical calculations, set up to counteract the lower solar luminosity, yield CO₂ atmospheric concentrations up to ten thousand times greater than today’s value of 0.035 %. Yet, this is at odds with the geologic record. For example, at low seawater pH, expected from such high partial pressures of carbon dioxide (pCO₂), ancient limestones should be enriched in 18O relative to their younger counterparts, yet the secular trend that we observe in the geologic record (Shields and Veizer, 2002) shows exactly the opposite. Factors more complex than a massive CO₂ greenhouse would have to be invoked to explain the warming of this planet to temperatures that may have surpassed those of the present day. A plausible alternative is a change in the cloud cover (Rossow et al., 1982) because clouds can compensate for 50% variations in radiative energy of the sun (Ou, 2001), bringing forward again the role of CRF as the potential solution. Considering that young stars of the same category as our sun would have been characterized by a stronger solar wind that muted the CRF, the resulting reduction in cloudiness may have compensated for the sun’s reduced luminosity (Shaviv, 2003). Note also that theoretical models of Milky Way evolution indicate a diminished star formation rate between ~ 2 and 1 Ga ago, while the Paleozoic and Neoproterozoic were strong maxima. This dovetails nicely with the geologic record (Frakes et al., 1992; Crowell, 1999), with massive glaciations at these two maxima and their absence in the intermediate time interval.

**CLIMATE ON MILLION YEAR TIME SCALES**

The record of climate variations during the Phanerozoic (Fig. 1) shows intervals of tens of millions of years duration characterized by predominantly colder or predominantly warmer episodes, called icehouses and greenhouse, respectively (Fig. 5). Superimposed on these are higher order climate oscillations, such as the episodic waning and waxing of ice sheets.

In the Phanerozoic, some organisms secreted their shells as the mineral calcite (CaCO₃), which often preserves

*Figure 4. Carbon isotopic composition of proxies for paleo-seawater, ancient limestones and calcareous shells (circles) and dolostones (triangles). Adapted from Shields and Veizer (2002).*
the original oxygen isotope ratio, and this, in turn, reflects the ambient seawater temperature. Veizer et al. (1999) generated a large database of several thousand well-preserved calcitic shells that cover this entire 545 million years time span. Such detrended isotope data correlate well with the climatic history of the planet (cf. Scotese, 2002; Boucot and Gray, 2001), with tropical sea surface temperatures fluctuating by perhaps 5 to 9°C between the apexes of icehouse and greenhouse times, respectively (Fig. 5, top).

The situation is entirely different for the CO2 scenario. For the Phanerozoic, the estimates of atmospheric pCO2 levels are not only internally inconsistent, but they also do not show any correlation with the paleoclimatic record (Fig. 5, bottom). In that case, what could be an alternative driving force of climate on geological time scales?

As suggested by theoretical considerations, the “icehouse” episodes and the oxygen isotope cold intervals should coincide with times of high cosmic ray flux, and the “greenhouse” ones with the low CRF (Fig. 6). This correlation may explain about 2/3 of the observed oxygen isotope “temperature” signal (Shaviv and Veizer, 2003). Thus celestial phenomena were likely the principal driver of climate on million year time scales.

CLIMATE ON MILLENNIAL TIME SCALES
Drilling at Vostok in Antarctica has produced an outstanding record of climate and atmospheric composition on millennial to centennial time scales for the last 420,000 years (Figs. 1, 7). The laminae of ice contain frozen air bubbles, and in these the amount of CO2 and methane indeed increases with temperature. Yet, new high-resolution studies show that at times of cold to warm transitions, temperature changes come first, leading CO2 changes by several centuries (Mudelsee, 2001; Clarke, 2003; Vakulenko et al., 2004). If so, the CO2 levels would be a response to, and not the cause of, the change in temperature (climate). CO2 may then serve as a temperature amplifier, but not as the climate driver.

If CO2 were not the driver, what could the alternative be? For the last 2 cycles of the Vostok record, spanning about 200,000 years, the residual geomagnetic field and the content of 10Be in sediments correlate antithetically (Fig. 8), at least at the 100,000 year frequency. 10Be is generated by the CRF interacting with our atmosphere. Since the solar and terrestrial magnetic fields are the shield that modulates the intensity of the CRF reaching the Earth, this anti-correlation is to be expected. The CRF, in turn, may regulate the terrestrial cloudiness and albedo, hence the climate. Having the estimates of the geomagnetic field intensity and 10Be concentrations enables calculation of the intensity of past solar irradiance. The latter appears to reflect surprisingly well coeval climate oscillations as recorded at Vostok and in the stacked oxygen isotope record of the oceans (Fig. 9). This points again to the

![Figure 5. Phanerozoic climatic indicators and reconstructed pCO2 levels.](image-url)
previously discussed extrinsic modulation of the CRF by the solar driver. Additional support for celestial forcing comes from ocean sediments and from caves, records that cover the times of transition from the last glacial episode into the warmer climates of our times, that is the time from about 11,500 to some 2,000 years BP. For an Atlantic drill core taken west of Ireland (Bond et al., 2001), the incidence of “ice rafted debris” (IRD), small debris pieces that fall to the ocean floor from melting ice floes that drift on the surface, coincides with the colder climates (Fig. 10). In addition, the cold times are characterized by high concentrations of $^{10}$Be, as measured in sediments, and by an “excess” of $^{14}$C, as observed in tree rings on land. Since both $^{10}$Be and $^{14}$C are products of the CRF interacting with our atmosphere, and because their subsequent redistribution pathways are entirely different, the only process that can explain all these positive correlations is an intensified CRF. Still better correlation is present in the cave sediments of Oman (Fig. 11). As stalagmites grow, they produce growth rings similar to those in the trees. The oxygen isotope ratio measured in these rings is a reflection of climate, in this particular case of monsoon patterns. The correlation with $^{14}$C, which is

**Figure 6.** The variations in the cosmic ray flux ($\Phi$) and tropical seawater temperature variations ($\Delta T$) over the Phanerozoic. The upper curves describe the reconstructed trends for cosmic ray flux (CRF) within their uncertainty band (stippled). The bottom curves depict the smoothed temperature anomaly (“Geological Reconstruction”) based on the $\delta^{18}$O record and the model cosmic ray flux (“Fit”). The peaks and valleys represent greenhouse and icehouse episodes as in Fig. 5. Note that no polar ice caps were as yet demonstrated for the third (hatched) icehouse. Adapted from Shaviv and Veizer (2003).

**Figure 7.** Antarctic (Vostok) ice core data for the last 400,000 years. Temperatures (dashed curve) are derived from oxygen and hydrogen isotopes of ice and CO$_2$ concentrations (dotted curve) were measured in frozen air bubbles. Adapted from Petit et al. (1999).
the product of CRF, is excellent. More recently, these cosmogenic nuclide/climate correlations were extended up to 2000 years BP and corroborated by additional records from an Alaskan lake (Hu et al., 2003), several European and American speleothems (references in Niggemann et al., 2003), polar ice shields (Laj et al., 2000; St-Onge et al., 2003), deep-sea sediments (Christl et al., 2003), and northern peat bogs (Blaauw et al., 2004) - geographic coverage of a considerable extent.

CLIMATE ON TIME SCALES OF CENTURIES

Let us now look at the record of the last millennium (Fig. 1), starting with Greenland, the climate record of the northern hemisphere. The calculations based on oxygen isotope values in ice layers suggest that the temperatures in the 11th century were similar to those of today (Fig. 12). This warm interval was followed by a temperature decline until the 14th century, then by generally cold temperatures that lasted until the 19th century, and finally by a warming in the 20th century. The “Medieval Climatic Optimum” (MCO) and the “Little Ice Age” (LIA), were both global phenomena (Soon and Baliunas, 2003; McIntire and McKitrick, 2003), and not, as previously claimed (Mann et al., 1999), restricted solely to Greenland or to the North Atlantic. Note that the coeval “ice bubble CO₂” pattern in Greenland and Antarctic ice caps was essentially flat (IPCC, 2001), despite these large climatic oscillations. CO₂ begins to rise only at the termination of the “Little Ice Age”, toward the end of the 19th century. In direct contrast to CO₂, 14C and 10Be correlate convincingly with the climate record (Fig. 13), again arguing for celestial phenomena as the primary climate driver.

THE DECADAL TO ANNUAL RECORD OF THE LAST CENTURY

The IPCC (2001) global mean surface temperature record shows an increase of about 0.6°C since the termination of the “Little Ice Age”. The bulk of this rise happened prior to the early 1940s, followed by a cooling trend until 1976 and a resumption of temperature rise subsequently (Fig. 14d). In contrast to temperature, the rise in atmospheric CO₂ most likely from the burning of fossil fuels

Figure 8. The intensity of the residual terrestrial magnetic field, and the 10Be content of marine sediments, for the last 200,000 years. Adapted from Sharma (2002). See also Christl et al. (2003).

Figure 9. Calculated intensity of solar irradiance (dots) during the past 200,000 years juxtaposed with the normalized δ¹⁸O record of the oceans (shading), the δ¹⁸O being a climate proxy. Adapted from Sharma (2002). In contrast to the CO₂/temperature correlation (Fig. 7), any potential causative sequence can only be from sun to earth, and temporal resolution is therefore not critical.
plus land-use changes, proceeded in an exponential fashion. This mismatch raises two questions: (1) why the large temperature rise prior to the early 40's, when 80% of the cumulative anthropogenic CO₂ input is post-World War II, and (2) why the subsequent three decade long cooling despite the rising CO₂? In contrast to CO₂, the temperature trend correlates well with the solar properties, such as the CRF and TSI (Figs. 14b,c), except perhaps for the last two decades of the 20th century that may or may not be an exception to this pattern. For these decades, the direct estimates of TSI flux (Fig. 14c) could not apparently explain the entire observed magnitude of the temperature rise (Ramaswamy et al., 2001; Solanki, 2002; Solanki et al., 2004; Foukal et al., 2004) and the discrepancy has to be attributed, therefore, to greenhouse gases, specifically CO₂. It is this discrepancy, and the apparent coherency of model predictions with observed climate trends (Karoly et al., 2003), that are the basis for the claim that the anthropogenic signal emerges from natural variability in the 1990's, with CO₂ becoming the “principal climate driver”. While this may be the case, note that the General Climate Models (GCMs) are essentially water-cycle models that generally do not incorporate the active carbon cycle and its dynamics. CO₂ is “prescribed” in most models as a spatially uniform concentration, and inputted in the form of energy (~ 4 Wm⁻² for CO₂ doubling). These models would yield outcomes in the same general direction, regardless of the source of this additional energy, be it CO₂ or TSI. Moreover, taking into account the empirical evidence, such as the unprecedented solar activity during the late 20th century (Fig. 13) or the coeval decline in global albedo (“earthshine”) (Fig. 15), and considering that the 1915-1999 TSI trend from the Mt. Wilson and Sacramento Peak Observatories can explain 80% of the 11-year smoothed variance in global temperature (Foukal, 2002), the celestial cause as a primary driver again appears to be a more consistent explanation. Additional support for such a scenario arises from the apparent relationship between solar cycle and precipitation/biological activity on land (Fig. 16). Terrestrial photosynthesis/respiration is the dominant flux for atmospheric CO₂ on annual to decadal time scales and any potential causative relationship can only be from the sun to the earth. As a final point, the GCMs predict that the most prominent centennial temperature rise should have been evident in the higher troposphere. Yet, the balloon and satellite data (Fig. 17) do not show any clear temporal temperature trend (IPCC, 2001). Instead, their interannual temperature oscillations correlate clearly with the solar irradiance and CRF, with “no vestiges of the anthropogenic signal” (Kärner, 2002). All this favours the proposition that celestial phenomena may have been the primary climate driver even for the most recent past.

In summary, the above empirical observations on all time scales point to celestial phenomena as the principal driver of climate, with greenhouse gases acting only as potential amplifiers. If solar activity accounts statistically for 80% of the centennial global temperature trend, while at the same time the measured variability in solar energy flux is insufficient to explain its magnitude, an amplifier that is causally related to solar energy flux should exist. The earlier discussed cloud/CRF link and/or UV related atmospheric dynamics could be such an amplifier(s). The existing general climate models may

Figure 10. The borehole record (Bond et al., 2001) of ice rafted debris (IRD), which is a climate proxy, and the coeval record for CRF proxies, ¹⁰Be in sediments and Δ¹⁴C in growth rings of trees. Adapted from Kromer et al. (2001).
therefore “require an improved understanding of possible climate sensitivity to relatively small total irradiance variations” (Foukal, 2002). I am aware that some of the discussed trends may have explanations based on the internal working of the earth system. For example, the $^{14}$C wiggles can be explained as changes in ocean circulation efficiency (ventilation), but this cannot explain the complementary $^{10}$Be patterns. In their sum, these explanations rely on many, at times arbitrary, causations and the overall structure is thus more complex than the celestial alternative. When two hypotheses can equally well explain the observational data, it is the simpler one that is to be preferred (Occam’s razor). I wish to emphasize, nevertheless, that it is not the intention of this contribution to discount superimposed geological, oceanographic, atmospheric and anthropogenic phenomena as contributing factors. Space considerations, however, do not allow this article to focus on anything but the nature of the “primary climate driver”.

**SO WHAT IS THE SEQUENCE?**
The review of empirical evidence strongly suggests that it may be the celestial phenomena, sun and cosmic rays, that are the principal climate driver. While the individual lines of evidence may have some weak points (but so do all alternative explanations), overall the celestial proposition yields a very consistent scenario for all time scales. The intrinsic CRF flux may have been responsible for the pronounced climatic trends on multimillion year time scales, while the extrinsic modulation by solar activity and earth dynamo could have been the major driver for the superimposed subdued climate oscillations on the millennial to annual time scales. This input drives the water cycle, with water vapour likely acting as a positive feedback and cloud formation as a negative one (Fig. 18). It also generates the flux of cosmogenic nuclides, such as $^{10}$Be, $^{14}$C and $^{36}$Cl. The hydrologic cycle, in turn, provides us with our climate, including its temperature component. On land, sunlight, temperature, and concomitant availability of water are the dominant controls of biological activity and thus of the rate of photosynthesis and respiration. In the oceans, the rise in temperature results in release of CO$_2$.
into air. These two processes together increase the flux of CO$_2$ into the atmosphere. If only short time scales are considered, such a sequence of events would be essentially opposite to that of the IPCC scenario, which drives the models from the bottom up, by assuming that CO$_2$ is the principal climate driver and that variations in celestial input are of subordinate or negligible impact. This is not to dismiss CO$_2$ as a greenhouse gas with no warming effect at all, but only to point out that CO$_2$ plays mostly a supporting role in the orchestra of nature that has a celestial conductor and the water cycle as its first fiddle. Consider an example that is familiar to every geologist, the weathering of rocks. This process is believed to have been the controlling sink for atmospheric CO$_2$ on geological time scales (Berner, 2003), and indeed it was. Yet, in reality, it is the water that is the agent of physical and chemical weathering. Weathering would proceed without CO$_2$, albeit with some chemical reactions modified, but not without water, whatever the CO$_2$ levels. For almost any process, and time scale, the water and carbon cycles are coupled, but water is orders of magnitude more abundant. The global water cycle is therefore not “just there” to react on impulses from the carbon cycle, but is actively shaping it. The tiny carbon cycle is piggybacking on the huge water cycle (clouds included), not driving it. In such a perspective, CO$_2$ can amplify or modulate natural climatic trends, but it is not likely to be their principal “driver”. If so, how are the global water and carbon cycles coupled?

**COUPLING OF THE WATER AND CARBON CYCLES**

The atmosphere today contains ~ 730 PgC (1 PgC = 10$^{15}$ g of carbon) as CO$_2$ (Fig. 19). Gross primary productivity

---

**Figure 13.** Time series of the sunspot numbers (reconstructed from $^{10}$Be in ice cores from Antarctica and Greenland), and of direct observations of sunspot numbers since 1610. The record of $^{14}$C in tree rings, not reproduced here due to visual considerations, shows a similar pattern. Note the low $^{10}$Be (reciprocal of sunspot numbers) and $^{14}$C during the Medieval Climate Optimum (MCO) and their high values during the Little Ice Age. Note also the very high solar activity for the latest 60 years, unprecedented for the last 8,000 years of Earth history (Solanki et al., 2004). W, S, M and D are the sun activity minima as in Fig. 12. O is the Oort minimum. Modified from Usoskin et al. (2003).

**Figure 14.** Eleven year average of northern hemispheric temperatures (dotted curve) and (a) solar cycle length (diamonds), (b) cosmic ray flux from ion chambers (dashed curve) and from the Climax neutron monitor (full curve), (c) solar irradiance (diamonds) (all modified after Svensmark, 1998), and (d) atmospheric CO$_2$ concentrations and global temperature trend (modified from Berner and Streif, 2000). Note that the polarity of CRF is reversed to facilitate comparisons and that the time frame of (a) to (c) represents only the post-1940 trough in (d).
(GPP) on land, and the complementary respiration flux of opposite sign, each account annually for ~ 120 Pg. The air/sea exchange flux, in part biologically mediated, accounts for an additional ~ 90 Pg per year. Biological processes are therefore clearly the most important controls of atmospheric CO2 levels, with an equivalent of the entire atmospheric CO2 budget absorbed and released by the biosphere every few years. The terrestrial biosphere thus appears to have been the dominant interactive reservoir, at least on the annual to decadal time scales, with oceans likely taking over on centennial to millennial time scales.

Interannual variations in atmospheric CO2 levels mimic the Net Primary Productivity (NPP) trends of land plants, and the simulated NPP, in turn, correlates with the amount of precipitation (Nemani et al., 2002, 2003; Huxman et al., 2004) (Fig. 16). The question therefore arises: is the terrestrial water cycle and NPP driven by atmospheric CO2 (CO2 fertilization) or is it the other way around? As a first observation, note that the “troughs” in precipitation and NPP coincide with the minima in sunspot activity (Fig. 16). If the sun is the driver, what might be the limiting variable? Except locally, CO2 cannot be this limiting factor because its concentration is globally almost uniform, while NPP varies by orders of magnitude. Temperature, because of its quasi anticorrelation with the NPP (Fig. 16), is not a viable alternative either. In contrast, the positive correlation between NPP and precipitation is clear-cut (Fig. 20) and water availability is therefore the first order limiting factor of ecosystem productivity (Huxman et al., 2004). Transpiration by ecosystems of cold and temperate regions recycles about 1/2 to 2/3 of precipitation into the atmosphere, while for tropical regions the recycling is almost wholesale. Thus the former appear to have been water starved (Fig. 20), while the tropical ecosystems with their efficient water recycling are likely limited only by the amount of available sunlight, the latter modified within relatively narrow limits, mostly by clouds.

For the global ecosystem, an increase in sunlight, humidity and temperature is a precondition for, not a consequence of, CO2 or nitrogen “fertilization”. And luckily so, otherwise our tree planting effort to sequester CO2 would only lead to a continuous massive pumping of water vapour, a potent greenhouse gas, from the soils to the atmosphere.

During photosynthesis, a plant has to exhale (transpire) almost one thousand molecules of water for every single molecule of CO2 that it absorbs. This so-called “Water Use Efficiency” (WUE), is somewhat variable, depending on the photosynthetic pathway employed by the plant and on the temporal interval under consideration, but in any case, it is in the hundreds to one range (Taiz and Zieglar, 1991; Telmer and Veizer, 2000). The relationship between WUE and NPP deserves a more detailed consideration. In plant photosynthesis, water loss and CO2 uptake are coupled processes (Nobel, 1999), as both occur through the same passages (stomata). The WUE is determined by a complicated operation that maximizes CO2 uptake while minimizing water loss. Consequently, the regulating factor for WUE, and the productivity of plants, could be either the atmospheric CO2 concentration or water availability. From a global perspective, the amount of photosynthetically available soil water, relative to the amount of atmospheric CO2, is about 250:1, much less than the WUE demand of the dominant plants, suggesting that the terrestrial ecosystem is in a state of water deficiency (Lee and Veizer, 2003).

The importance of the water supply for plant productivity is clearly evident from the NPP database that is a collection of worldwide multi-biome productivities, mostly established by biological methods (Fig. 20). The principal driving force of photosynthesis is unquestionably the energy provided by the sun, with the global terrestrial system reaching light saturation at about an NPP of 1150 ± 100 g carbon per year (Fig. 20). If the sun is the driver, what might be the limiting variable? Except locally, CO2 cannot be this limiting factor because its concentration is globally almost uniform, while NPP varies by orders of magnitude. Temperature, because of its quasi anticorrelation with the NPP (Fig. 16), is not a viable alternative either. In contrast, the positive correlation between NPP and precipitation is clear-cut (Fig. 20) and water availability is therefore the first order limiting factor of ecosystem productivity (Huxman et al., 2004). Transpiration by ecosystems of cold and temperate regions recycles about 1/2 to 2/3 of precipitation into the atmosphere, while for tropical regions the recycling is almost wholesale. Thus the former appear to have been water starved (Fig. 20), while the tropical ecosystems with their efficient water recycling are likely limited only by the amount of available sunlight, the latter modified within relatively narrow limits, mostly by clouds.

Figure 15. Reconstructed annual reflectance anomalies (Δp*) relative to 1999-2001 calibration interval (shaded). The observed anomalies are represented as a thick line. In general, Δp* is a measure of earth albedo, likely cloudiness, by observing the "earthshine", the light reflected by Earth’s sunlit hemisphere toward the moon and then retroflected from the lunar surface. Note that the decline in albedo (cloudiness) from 1985 to 2000 is a feature that is consistent with the increase in solar irradiance TSI (Fig. 13) and implicitly also with a decline in cloud nucleation due to diminished CRF. Note also that the cloud-driven changes in the Earth’s radiation budget (up to 10 Wm⁻²) during the last two decades exceed considerably the forcing that is attributed by IPCC (2001) to the entire “industrial”, that is post "Little Ice Age", anthropogenic greenhouse impact (2.4 Wm⁻²). Adapted from Pallé et al. (2004a).
In order to test the hypothesis of CO₂ “piggybacking” on the water cycle, several large watersheds were examined, because there the water balance can be deconvolved into precipitation, discharge, evaporation, interception and transpiration fluxes. Knowing the transpiration flux and the requisite WUE, it is then possible to calculate the photosynthetic sequestration capacity for CO₂ for a given watershed. Taking the Mississippi basin (Fig. 20) as an example (Lee and Veizer, 2003), plant transpiration recycles about 60% of precipitation back into the atmosphere and the calculated, water balance-based, annual photosynthetic sequestration of CO₂ by plants is then 1.16 Pg of carbon. This is essentially identical to the heterotrophic soil respiration flux of 1.12 PgC derived by biological approaches for the same watershed. Hence, the suggestion that the carbon cycle is “piggybacking” on the water cycle is a viable proposition. This scenario is supported also by the satellite data of global productivity for the 1982-1999 period, with “climatic variability overland exerting a strong control over the variations in atmospheric CO₂” (Nemani et al., 2003). In these two decades the global biomass grew by 6% (3.4 PgC). Almost one half of the increase happened, surprisingly, in the

Figure 16. 1900-1993 variations in annual averages of air temperature (T - dotted line) and precipitation (P - dashed line) for conterminous U.S. together with the simulated Net Primary Productivity (NPP - full line) smoothed with a 5-year filter (adapted from Nemani et al., 2002). The arrows are the years of sunspot minima (dampened solar irradiance) from the Royal Observatory of Belgium (http://sidc.oma.be/index.php3). Note that except for 1944, the troughs in precipitation and NPP appear to coincide with the sunspot minima. Figure courtesy of Ajaz Karim.

Figure 17. Annual variability in tropospheric temperature, TSI (ΔFₚ) and CRF. Note the reversed scale for CRF. Modified from Marsh and Svensmark (2003b) and Marsh et al. (2005).

Figure 18. Schematic presentation of the sequence of events for a model based on celestial forcing as the principal climate driver. The dashed arrow is a feedback from the biosphere on climate, including its anthropogenic component.
Amazon basin, and was caused by a decrease in the cloud cover (decline in CRF?) and to a concomitant 20th century increase in solar radiation (Figs. 13, 14, 15). Again, while CO₂ may act as an amplifying greenhouse gas, the actual atmospheric CO₂ concentrations are controlled in the first instance by the climate, that is by the sun-driven water cycle, and not the other way around.

ENVIRONMENTAL IMPLICATIONS

At this stage, two scenarios of potential human impact on climate appear feasible: (1) the standard IPCC model that advocates the leading role of greenhouse gases, particularly of CO₂, and (2) the alternative model that argues for celestial phenomena as the principal climate driver. The two scenarios are likely not even mutually exclusive, but a prioritization may result in different relative impact. Models and empirical observations are both indispensable tools of science, yet when discrepancies arise, observations should carry greater weight than theory. If so, the multitude of empirical observations favours celestial phenomena as the most important driver of terrestrial climate on most time scales, but time will be the final judge. Should the celestial alternative prevail, the chain of reasoning for potential human impact may deviate from that of the standard IPCC model, because the strongest impact may be indirect, via the formation of cloud condensation nuclei (CCN). The CRF-generated positive and negative ions combine, within minutes, into electrically neutral aerosols, but only if the two ions are large enough. The required size of these “cluster ions” is reached by addition of atmospheric molecules, particularly sulphuric acid. Since H₂SO₄ is highly hygroscopic, it attracts also water molecules. In this way, the ~30-100 nm large CCN required as precursors for droplets can potentially be generated (Carslaw et al., 2002; Lee et al., 2003). Thus, sulphur compounds (and perhaps dust, soot and secondary particles, which are formed by condensation of low-vapour-pressure gases) could play a major role in this seeding process. In the northern hemisphere, the precursor of sulphuric acid, sulphur dioxide gas, originates mostly from anthropogenic activities, but natural sources, such as volcanic eruptions or DMS from marine plankton, are also substantial.
Although the role of clouds is not well understood (IPCC, 2001), it appears that the upper tropospheric clouds warm, while the lower clouds, such as those potentially generated by the above CRF seeded processes, cool the climate. In such a scenario, the impact of pollution, if indeed significant, could even potentially result in global cooling (Carslaw et al., 2002) instead of global warming, similar to the IPCC chain of reasoning that is invoked as an explanation for the 1940-1976 cooling trend (Fig. 14d). In addition, we would have to deal not with a global issue of atmospheric CO2, but with large regional phenomena, because it is these that control the dispersion of aerosols, sulphur and nitrogen compounds. We are not yet in a situation where quantitative projections of this impact on climate can be provided (Schwartz, 2004). Indeed, we do not even know if it is at all globally significant, equal to any potential warming generated by CO2, or much larger. In any case, the strategy that emphasizes reduction of human emissions is sound because it is these that control the dispersion of aerosols, sulphur and nitrogen compounds. These are not only potential climate drivers, but also pollutants whose reduction will improve our air quality, regardless of the climate impact of otherwise environmentally benign CO2. At current atmospheric levels, CO2 is in fact an essential commodity for propagation of life on this planet. Any remedial measures based on the global CO2 scenario are also costly. For the celestial alternative, the remedial measures may focus directly on the “collateral” pollutants, which could potentially result in a substantial reduction of the economic cost to mankind. However, the decision as to the best strategy is not a simple prerogative of science, but must also take into account political, economic and social considerations.

ACKNOWLEDGEMENTS
In my four decades of research into the evolution of the Earth, always with strong environmental connotations, I was almost exclusively financed by the Natural Sciences and Engineering Research Council of Canada (NSERC) and the Deutsche Forschungsgemeinschaft (DFG). In the last decade, particularly relevant to this article, the research was supported by two major sources, the top research award of the DFG (Leibniz Prize endowed with 3 million DM) and the support of the Research Chair in “Earth System” financed jointly by NSERC and the Canadian Institute for Advanced Research (CIAR). The donors to CIAR include Noranda and Dr. G.G. Hatch, with the sponsorship based on an arms’ length relationship via CIAR and NSERC.

Personally, this last decade has been a trying period because of the years of internal struggle between what I wanted to believe and where the empirical record and its logic were leading me. This article is clearly not a comprehensive review of the alternatives, partly because of space limitations, but also because the case for the alternatives was eloquently argued elsewhere (e.g., IPCC, 2001). It is rather a plea for some reflection in our clamour for over-simplified beliefs and solutions in the face of the climate conundrum. Due to space considerations, the article also does not explore the potential role that the lethal CRF may have played in the evolution of life, as a cause of extinctions and/or mutations. And above all, this article is not a discussion of Kyoto, a treaty with social, economic and political aims, but a scientific treatise of the past climate record. Time will rule on its validity, but in the meantime I ask that the discussion of its merits/demerits be confined to scientific ways and means.

As a final point, I am indebted to several experts worldwide, covering the whole gamut of fields from astrophysics to biology and modeling, who agreed to read the manuscript in order to make sure that its statements are scientifically defensible. The journal reviewers, Brendan Murphy and Alan Hildebrand, helped to set the tone of the presentation.

REFERENCES


De la Fuente Marcos, R. and de la Fuente.


