Milankovic Pseudo-cycles Recorded in Sediments and Ice Cores Extracted by Singular Spectrum Analysis

Fernando Lopes[1], Pierpaolo Zuddas[2], Vincent Courtillot[1], Jean-Louis Le Mouël[1], Jean-Baptiste Boulé[3], Alexis Maineult[2], Marc Gèze[4]

Response to Reviewer 1

We acknowledge this review but find unwarranted criticism, and erroneous comments in this review, which we reject almost in its entirety.

We respond below (in black) to the reviewer's comments (in red, numbered), first in short summary form, then in more detail.

Short summary of the reviewer's criticisms and our responses:

1) Finding frequencies in a signal is NOT (has never been) the key point of spectral analysis. Finding many frequencies is easy and NOT the main problem of "spectral analysis".

Wrong in large part.

2) "Paleoclimatic records contain astronomical periodicities" vs " It is well known since the 19th century that global insolation changes are much too weak to affect climate".

Self contradiction.

3) All papers since Wegener consider only summer insolation at high northern latitude vs global averaged insolation.

Wrong. We have written a detailed analysis of Wegener (1920) as Note I below.

4) Spectral analysis is a branch of statistics. No discussion of the "detection level" problem. Embedding dimension too high.

All wrong. We have written a detailed analysis of SSA and statistics as Note II below.

5) Classical subject, brings nothing new. Authors not familiar with this topic.

Wrong.

Detailed responses to the reviewer's criticisms, including two more extensive notes, one on Milankovitch theory, the second on the SSA method.

1: I have difficulties to understand the main topic of this manuscript: Either the main message is that paleoclimatic records contain astronomical periodicities, something which is well-known for almost 50 years. Or, more probably, the aim is to present a "new" spectral analysis method that could detect extremely faint periodicities in geological records. In the first case, I would reject the paper on the ground that it brings nothing new.

Of course the paleoclimatic records contain astronomical periodicities; that has been known at least since Milankovitch (1920). But the question here is to use SSA to determine how many of these periodicities (SSA components) are present and what are their respective amplitudes (contributions to the original signal). And we succeed in doing it and our results are indeed in part new (number, periodicity content and amplitude (contribution to the total variance) of main contributing (quasi-)periodicities (components) determined using SSA, more complete than found in previous publications to our knowledge). See also Note I.

2: But the authors are not even citing the famous foundational Hays et al (1976) paper in their reference list, which is, to say the least, very peculiar and which suggest that they may not be familiar with this topic and its extremely abundant literature.

We do know and cite Hays et al (1976) in our text, but we have indeed forgotten it in the reference list. We will correct this upon revision. We note that that paper only finds 7 periodicities (in δO^{18}) when we identify 5 more with periods longer than 40 kyr (in both CH4 and CO2 in the same Vostok core): 50, 66, 100, 180 kyr. Our reference list and previous published papers show that we have some familiarity with the topics involved in our paper.

3: Besides (see below), it appears that the authors have even missed the main point of Milankovitch theory, which is based on summer insolation, and not global mean insolation...

Wrong. Note I below answers this and other comments at some length.

NOTE I (ON MILANKOVIC THEORY)

Milankovich's thesis (1920, in French) :
https://www.dropbox.com/s/snv3fdc601b55ws/Milankovic1920.pdf?dl=0

Milankovich (1948) (in German) :
https://www.dropbox.com/s/67z2j5hcrbrdhsw/milankovitch1948.pdf?dl=0

We refer to Milankovitch's original papers (that are in French; f.i. Part II, chapter 1, page 169, entitled « Distribution de la radiation solaire à la surface du globe terrestre et son climat mathématique »). In short, Milankovitch does not mention summer insolation but global insolation. We now follow in detail Milankovitch's reasoning.

The last sentence of the reviewer is indeed encountered in many if not most papers on present and past climate. We claim that it is both wrong and rather disconcerting.

Why disconcerting? Let us use the case of marine sediments from Lisiecki and Raymo (2005); their time sampling, which is not regular, has a mean of 2500 years. What is then the physical meaning of considering only summer, *i.e.* 3 months of the year, when the time series under analysis has a sampling rate of several thousand years and the goal is to determine astronomical parameters the smallest of which is 20,000 years ?

Why wrong? Let us go back to the writings of Milankovitch (1920, 1948). We give a link to the full texts, the former in French, the latter in English and German.

The latter is not scientifically innovating with respect to the former as he himself states (left column, page 414, second paragraph);

« *In dem Buche Théorie mathématique des phénomènes thermiques produits par la radiation solaire (Paris 1920) erscheinen die wichtigsten Teile des gestellten Problems in ihren Hauptziigen bereits als ge1öst, wodurch das begonnene Lehrgebäude seine festen Grundmauern erhielt, auf die man weiter bauen konnte* ». Milankovitch (1948), written 8 years after Koppen's death, tells how he, Koppen, and his son in law Wegener needed help from mathematician Milankovitch for some computations, those same computations that had led Milankovitch to the axonometric Figure 13 of his 1920 thesis (part 2, page 182), representing the variations of the diurnal insolation as a function of geographic latitude (-90° to +90°) for all solar longitudes (0 to 360°). This is clear from the following paragraph (Milankovitch, 1948, page 414, right column, end of paragraph: « *So wurden also, dem Wunsche KÖPPENS entsprechend, jene Änderungen berechnet, die die sommerliche*

Bestrahlung der geographischen Breiten von 55, 60 und 65° nördlich während des Intervalls der letztverflossenen 650 Jahrtausende erfahren hat, wobei ich mich darauf beschränkte, die Amplituden dieser Änderungen zu berechnen und in Zackenkurven darzustellen. »

Milankovitch's (1920) thesis is made of two parts, one theoretical, the other more applied. From here on, all equations and page numbers are those of the thesis. Milankovitch shows that insolation at a given point of the Earth's surface (with latitude φ and longitude ψ), at time *t* is (page 15, relation 22):

$$\frac{dW}{dt} = \frac{I_0}{\rho^2}(\sin\varphi\sin\delta + \cos\varphi\cos\delta\cos\psi_1)$$

with I_0 being the solar constant, ρ the Earth-Sun distance, δ the Sun's declination and ψ_1 the sum of the point's longitude and the solar right ascension (ω). In the course of a given day, ρ and δ can be considered as constants, and one can write (relation 32, page 19) :

$$\frac{dW}{dt} = A + B\cos\omega$$

where A and B are constants for the given point. This relation holds for values of ω that yield a positive value for inclination, that is between the hourly angles of sunrise and sunset. These angles are given by the equation (relation 33, page 19):

$$\cos\omega_0 = -\frac{A}{B} = \tan\varphi\tan\delta$$

Root $-\omega_0$ corresponds to sunrise and root $+\omega_0$ to sunset. When the following inequality (relation 39, page 21) holds,

$$-1 < \tan\varphi\tan\delta < 1$$

ω_0 is real and the Sun disappears under the horizon during a time interval shorter than the day. Since the Sun's declination δ cannot be larger than and smaller than $+\varepsilon$ (ε being the inclination of the axis of rotation, i.e. of the pole), it comes that the preceding inequality holds all along the year for geographical latitudes between:

$$-(\frac{\pi}{2} - \varepsilon) < \varphi < (\frac{\pi}{2} - \varepsilon)$$

with present values between -66.55° and +66.55°. The parallels at $\varphi = \frac{\pi}{2} - \varepsilon$ and $\varphi = -(\frac{\pi}{2} - \varepsilon)$ are the polar circles. In that zone, the Sun rises and sets every day. At latitudes $\varphi > \frac{\pi}{2} - \varepsilon$ and $\varphi < -(\frac{\pi}{2} - \varepsilon)$ the daily march of the Sun is different. Milankovitch develops the mean daily insulations for different latitudes. We skip this and come to the second parameter that is important to calculate the irradiation that is received, namely the Solar longitude (λ). On page 37 he gives the different astronomical seasons as a function of λ ($\lambda = 0$ for the spring equinox, $\lambda = \frac{\pi}{2}$ for the summer solstice, $\lambda = \pi$ for the autumn equinox and $\lambda = \frac{3\pi}{2}$for the winter solstice). He shows on pages 37 to 40 that his system (91) of equations, independently from latitudes (for both hemispheres), implies that (page 39): «*Celles-ci nous montrent que la quantité de radiation correspondant à l'été est égale à celle du printemps, tandis que celle de l'hiver est égale à celle de l'automne.*». Milankovitch calls W_I the amount of irradiation received during Spring for λ between and 0 in $\pi/2$, the boreal hemisphere (page 37), W_{II} the amount received in Summer, still in the boreal hemisphere, for λ between $\pi/2$ and π, W_{III} the amount of irradiation received during Autumn for λ between π and $3\pi/2$ in the boreal hemisphere and W_{IV} the amount of irradiation received during Winter for λ between $3\pi/2$ and 2π. In the same way, he symmetrically defines for the austral hemisphere the following quantities: $\bar{W}_I$, $\bar{W}_{II}$, $\bar{W}_{III}$ and $\bar{W}_{IV}$. He concludes (system 91, page 39) that $W_I = W_{II} = \overline{W}_I$. He defines $W_e = 2W_I$ and $W_h = 2W_{III}$, W_e and W_h being respectively the amounts of heat received during the warm and during the cold periods, for a given latitude. He then describes the behavior of W_e and W_h all the way to the end of part I (page 165). Thus, where does the reviewer find the original references that allow him/her to conclude, falsely, that "the authors have even missed the main point of Milankovitch theory, which is based on summer insolation, and not global mean insolation?"

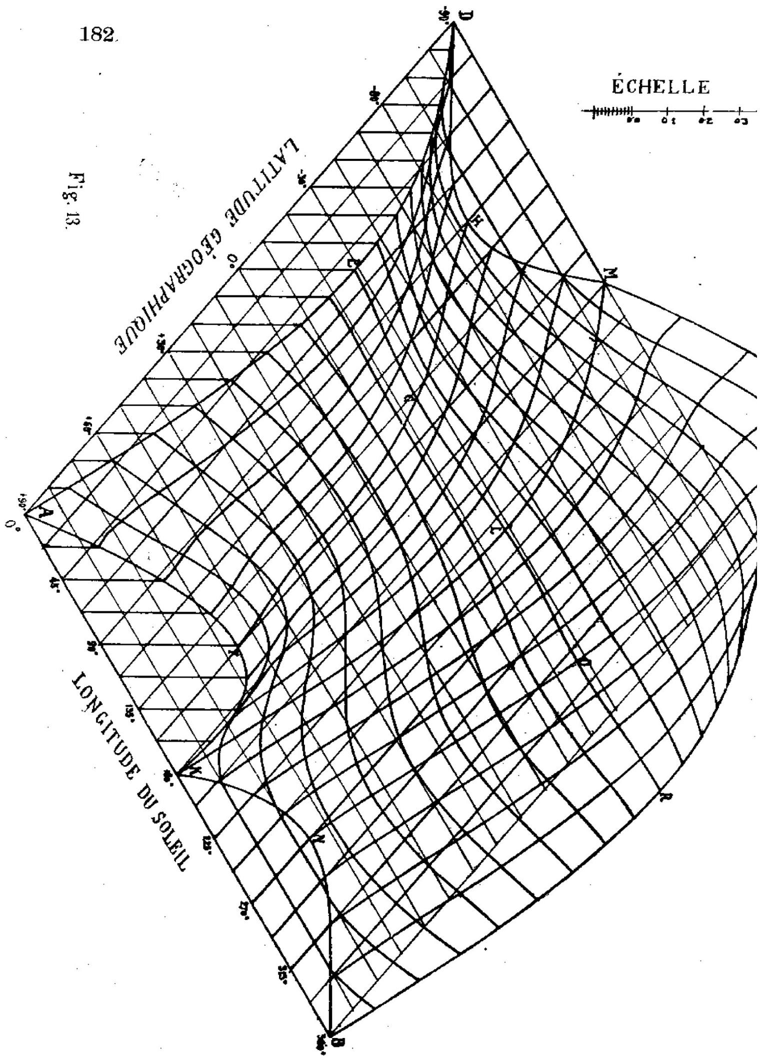
We now come to the more applied part 2 of Milankovitch's (1920) thesis. Milankovitch first presents results obtained by other authors regarding the amount of daily variation of insolation as a function of the Sun's longitude for ALL geographical latitudes (-90° to +90°). Tables II, pages 176-177, come from Wiener's theory. The amount of daily radiation transmitted to and received by the ground surface, for different values of the transmission coefficient for solar radiation as a function of the axis inclination, here again for latitudes going from the equator to the North pole (Tables III, pages 179-180) is from the work of Angot. Milankovitch summarizes these first results on page 181 by the following relation for $\lambda = 0$, that is to say the Spring equinox for zero solar declination δ and unit Solar constant:

$$W_\tau = w = \frac{1}{\pi \rho^2} \cos \varphi$$

« *A ce moment la distribution méridionale de la quantité diurne de radiation et de la moyenne insolation du parallèle suit une loi simple: les quantités sont proportionnelles au cosinus de la latitude géographique, atteignent leur maximum à l'équateur pour disparaître aux deux pôles.*» In regions where the Sun doesn't set, Milankovitch shows that the daily variation is given by:

$$W_\tau = \frac{\sin\varphi \sin\epsilon \sin\lambda}{\rho^2}$$

Milankovitch represents these two equations for the mean radiation as an axonometric figure (Figure 13 page 182) that we reproduce below.



It is therefore quite clear that Milankovitch does take into account all latitudes. It is this very same Figure 13 that is referred to as *Zackenkurven* in Milankovitch (1948).

We next come to the topic that Milankovitch calls "*Variations séculaires des éléments* *astronomiques et le problème paléoclimatique*" (page 221, Chapter 2). We attempt to follow Milankovitch's reasoning in his thesis as closely as possible. At the beginning of page 221, the author recalls the 4 parameters that influence the Earth's insolation: the eccentricity of the planet's orbit, the longitude of the perihelion, the general precession and the inclination of the planet's axis of rotation.

Thanks to formulas derived by Stockwell and to his own equations (12) and (13) (pages 12 and 13), Milankovitch makes a numerical calculation of the consequences of increasing obliquity by 1° on amounts of total insolation (W_T), for the warm period (W_e) and the cold period (W_h), for all latitudes from 0° to 90°: see Tables XVII and XVIII on page 228. There is NO mention of a limitation to certain latitudes or (summer?) season. On the contrary, Milankovitch writes about variations in percentage of insolations on page 229: "*Elles présentent le fait remarquable que les quantités de variations hivernales correspondant aux latitudes 60 à 70 degrés sont très influencées par les changements d'obliquité de l'écliptique"* Milankovitch next calculates these variations at key times of the Earth's revolution, that is solstices and equinoxes, leading to the systems of equations on pages 231 et 232 (in part 2 of his 1920 thesis/book, Milankovitch does not give numbers for equations any more). He writes on page 231: " *Ces équations sont valables pour la zone non-arctique* [note: between $\pi/2 - \varepsilon$ and $-(\pi/2 - \varepsilon)$], *tandis que pour les zones arctiques on devrait employer les formules simples (38) dont nous ne nous occuperons plus dans ce qui suit"*. Here again there is no limitation in the range of latitudes in Milankovitch's theory.

Milankovitch's study of the secular variation of insolation leads to Figure 20 page 239, which represents the famous 40,000 year cycle, imposed by variations of and he calculates the solutions for W_e and W_h. Milankovitch also demonstrates that (page 244, end of first paragraph): " *Il découle de ce qui précède que l'atmosphère renforce les effets des variations de l'obliquité de l'écliptique"*. This means that if the rotation axis "straightens up", then it will be colder in the Summer (the Earth being farther from the Sun) and warmer in Winter (the Earth being closer to the Sun), as is indeed observed currently.

We now come to chapter 56, part 2 (page 246), the most relevant to our study and to this response to the reviewer, entitled "*Théories astronomiques des époques glaciaires"*. Milankovitch first presents his master Adhémar's theory, that he writes is "*sous l'empire des idées de Fourier"* (page 247) and that he finally rejects. So does he reject Croll's theory, of which he acknowledges only the idea of an astronomical forcing. In this rejection, there is a remarkably simple development, grounded on observations. For instance, he writes on page 251 that beyond the 66th parallel ($\pi/2 - \varepsilon$), a zone that in Milankovitch's theory obeys a second set of equations, "*Par conséquent à ces latitudes, on peut obtenir presque les mêmes effets par la variation de l'obliquité* *de l'écliptique que par la variation de la durée des saisons".*

This is an important remark, as soon as one is concerned by geochronology at these latitudes. This is seen in the data as he clearly spells out on page 252: "*Ceci ressort également de la Figure 21, d'après laquelle l'insolation hivernale* [note: NOT the Summer insolation] *du 48^e degré de latitude Nord* [note: NOT beyond the polar circle] *n'exécute, dans l'intervalle de la 72^e à la 25^e milliade avant l'époque actuelle, qu'une seule oscillation, tandis que , par la suite de la variabilité des saisons, elle en subirait deux".* Behind this cogent remark lie many problems in geochronology and paleoclimatology. Milankovitch next criticizes Ball's (page 253), Eklom's and Spitaler's (page 254), only to finally apply his own theory.

He starts by recalling that: "*Pour l'étude du climat du passé géologique, il importe de connaître les limites entre lesquelles le climat mathématique de la Terre peut varier par suite de la variabilité des éléments astronomiques. Ce cadre est tracé par les limites entre lesquelles peuvent varier les éléments astronomiques et qui ont été donnés au numéro 52".* Paragraph 52 of chapter II (Part 2, page 222) leads to a system of equations that evaluate the insolation received by Earth, for ALL latitudes and for warm AS WELL AS cold periods, as a function of time.

Milankovitch then only has to assign all astronomical variables their extreme values to calculate the range of the resulting radiation. The results are summarized in Table XXII (page 256) in which one sees that for all latitudes, the influence of obliquity of the ecliptic leads to extreme values of 24°36 and 21°58. He next evaluates (given the duration of current seasons (summers and winters) the maximum and minimum departures that this radiation can attain as a function of ALL latitudes. His Table XXIII (page 259) summarizes his computations. The next figure (number 24, page 262) in Milankovitch's thesis represents the values listed in Table XXIII. From there on, he deduces for instance that, for an extreme obliquity of 21°58 during the 186.46 days of strongest insolation, the 90° parallel receives as much insolation as today's 69° parallel, the 70° as much as the 60°, the 50° as the 30°, etc… (see the values under accolades on pages 260 et 261). This is one of the reasons that will lead him to consider polar drift, and also why it is somewhat strange to consider climatic phenomena only beyond the 65° parallel. Milankovitch next considers the question of extreme insolations (Part 58, page 263). He first recalls that " *Les moyennes températures annuelles ne dépendent que des quantités annuelles de radiation* W_T *et* W_T'*; c'est pourquoi ne sont-elles influencées que d'une façon insignifiante par la variabilité de la valeur de (excentricité) et pas du tout par celle de* Π *(longitude du périhélie par rapport au point vernal)* ".

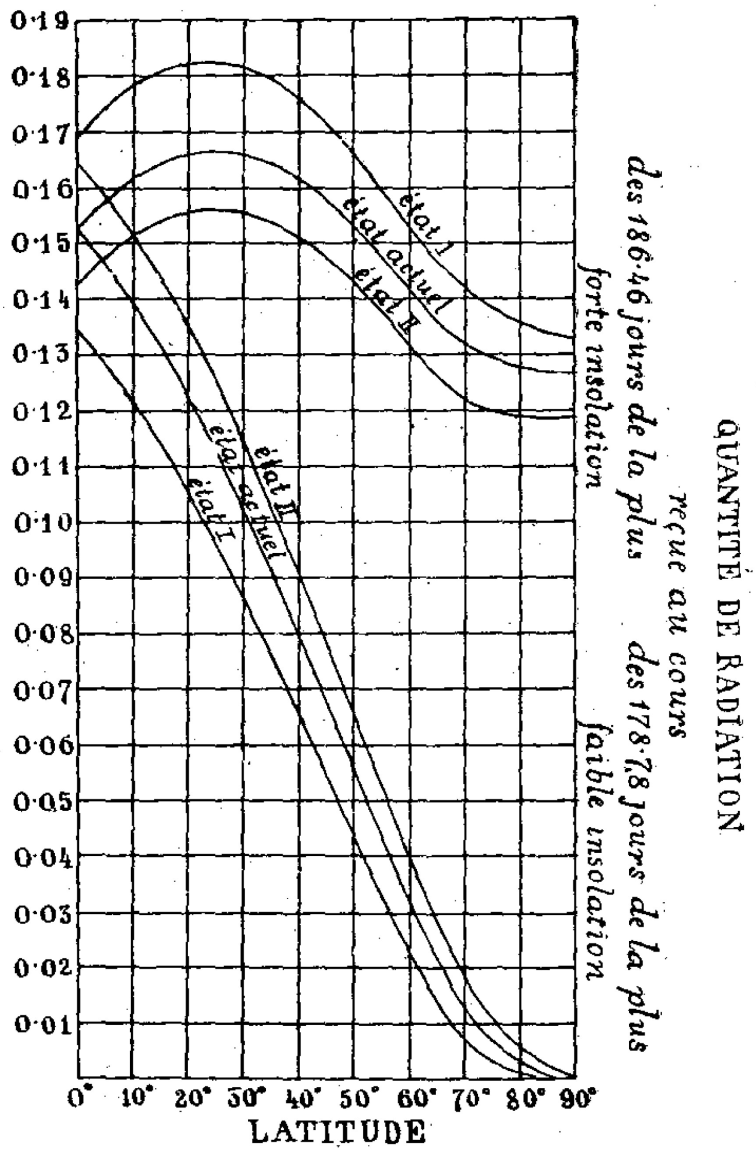


Fig. 24.

So, it is clear from his very own writing that temperatures are influenced by the mean total amount of radiation received, not by the summer amount. Milankovitch further specifies " *A cet effet, on doit calculer, d'abord les valeurs de* W_T *et* W'_T *pour les deux cas extrêmes d'insolation. Les valeurs de sont aux pôles, égales aux valeur de* W_e*"*. Indeed, there is no daily alternation of insolation at the poles and the maximum insolation is de facto W_e that is the insolation of the warm period (Spring + Summer), which is equal to the total insolation in the year. After integrating the effect of atmospheric absorption, Milankovitch deduces that for extreme values of obliquity (page 265) "*En comparant ces nombres à la valeur de -34.8°, donnée du numéro 48, on voit que la température annuelle solaire du pôle peut augmenter de 3.2° et diminuer de 4.2°"*. And these values are constants, since there is no alternation between day and night at the poles. After having dealt with the Arctic zone, Milankovitch next addresses the non-Arctic zone "*En s'éloignant des pôles,* *ces oscillations de la température annuelle diminuent rapidement pour devenir à peine notables dans la zone modérée* [note: tropical and equatorial]. *La détermination de ces oscillations possibles est très difficile, le calcul de étant une opération laborieuse. Aussi n'allons-nous plus nous en occuper ".* However, after several simplifying assumptions, such as noting the weak eccentricity of the Earth's revolution, Milankovitch computes the extreme variations in amplitudes of the annual temperature (still with the same two values of obliquity) for latitudes -65° to -25° and 25° to 65°. This is summarized in Table XXIV, page 268. Thus, he covers insolation in a precise way in the polar regions, in a simplified way elsewhere.

In parts 59 and 60, Milankovitch looks into the other causes that could lead to climatic modifications, and foremost polar drift. He cites the work of Darwin (page 270), noting that the polar caps play the role of a "*patin à frein*", as is the case for the tidal breaking due to the Moon's attraction. He also cites Chandler (pages 271 to 273) and the effect at very long periods named after him. He ends with Wegener's theory coupled to the classical theory of an elastic Earth, under which a reorganization of continental masses at the Earth's surface could lead over very long periods to modifications of the rotation axis (polar drift). Clearly, Milankovitch followed Stockwell and not either Laplace, or Poincaré. He also mentioned variations in atmospheric composition (thus of its absorption) and in the solar constant I_0 . But as he himself states, all these remain hypothetical.

Milankovitch concludes the application of his theory to the Earth by paragraph 61 (page 279) with the title "*le problème paléoclimatique"*. In the introduction of this chapter, Milankovitch first warns the reader " *...car si nous voulions poursuivre la marche séculaire de l'insolation et du climat mathématique aussi loin que nous le permettent les lois de la Mécanique céleste, nous serions obligés de mettre en jeu un matériel numérique immense dont il n'y aurait probablement qu'une minime partie qui pourrait être utilisée dans les recherches ultérieures "* (page 280).

So, one can forget about a theory of paleoclimates based on computations alone beyond the 65° parallel. Milankovitch goes on (top of page 281) " *Le but de cet ouvrage est fournir l'outil nécessaire pour des recherches de ce genre et non point de s'étendre sur ces recherches mêmes. Nous ne ferons que signaler où elles pourraient être employées avec succès "*.

This entire paragraph is an attempt by Milankovitch to explain climatic variations observed with geology and archeology in Europe (mainly in Northern Europe) and in North America during the Pleistocene.

In conclusion of this section, we have attempted to show (or rather recall) that Milankovitch (1948) is not so much an innovative scientific paper but rather an acknowledgement of Wegener and Koppens in which Milankovitch recalls their joint collaboration and in which he specifies topics that had all been fully treated in his (1920) thesis. Milankovitch himself performed the computations of summer insolation at latitudes 55°, 60° et 65° upon a request from Koppens himself. Koppens had certainly not been able to perform them without help. In his 1920 thesis, that we have followed page after page in this answer to our reviewer, never did Milankovitch write that his theory did not function beyond 65°. On the contrary, he generalized his work to show that daily alternations of sunrise and sunset occur between latitudes $-(\pi/2-\varepsilon)$ and $\pi/2-\varepsilon$. He developed a system of equations for this band of latitudes. He developed another system for complementary bands, that is the polar regions. Both systems were summarized in his Figure 13 (reproduced above). As far as the summer variation is concerned, Milankovitch never takes it into account, except to give some examples. On the contrary, he divides the insolation into two parts: the warm periods (W_e), corresponding to the spring/summer couple and the cold periods (W_h), corresponding to the autumn/winter couple. As far as glacier life is concerned, Milankovitch says that it remains a hypothetical story (Paragraph 61, page 279), for things are very complex. Altogether, Milankovitch explicitly writes that polar insolation during warm periods (W_e) is equal to total mean annual insolation (W_t).

END OF NOTE I

4: In the second case, I would also reject the paper since it does not show any information on the key question of spectral analysis: statistical relevance.

We have provided the key reference for all practical and mathematical aspects of SSA (Golyandina, N., and Zhigljavsky, A.,: Singular Spectrum Analysis for time series, Vol. 120, Berlin, Springer, 2013), in particular statistical relevance. We do give uncertainties for all component periods in Table 1. We devote Note II to SSA and statistics to refute the reviewer's criticism.

This note recalls two methods, one of which (SSA) was used in this paper. Its aim is to recall the bases of frequency analysis, and to compare them to SSA. This note is a summary, but a rather extensive one, given the fact that SSA does not seem to have enjoyed the success one could have imagined. Also, this is to make sure that the reviewer shares our understanding and competence in using the method.

NOTE II (ON SPECTRAL ANALYSIS , STATISTICAL ANALYSIS AND SSA)

There are several definitions of spectral analysis (e.g. Claerbout, 1976 or Bracewell and Bracewell, 1986). We choose the simplest and most basic one (*cf.* Stoica and Moses, 2005, chapter 1.1 page 1): "*From a finite record of stationary data sequence, estimate how the total power is distributed over frequency*". Spectral analysis, the most ancient branch of harmonic analysis, of whom Fourier series (Prony, 1795, Fourier, 1822) are the most ancient representative, consists in projecting any stationary (in a strict sense, statistically) signal on an orthogonal basis of sine functions. Its limitations are well known (*e.g.* Shannon and Burks, 1951; Kay and Marple, 1981). The discrete Fourier transform (S_τ) of any time series $s(t)$ sampled every τ and represented by N samples, is given by the following equation :

$$S_{\tau,T}(k\nu) = \tau \sum_{n=0}^{N-1} s(n\tau) W^{kn} \qquad (k = 0, 1, \ldots, N-1) \tag{01}$$

The sampling rate ν in the frequency domain must satisfy Shannon's dual law ($\nu < \frac{1}{T}$) where T is the duration of the signal. k is an integer describing the position of a given frequency in the dual space, whose maximum value must be at least equal to N (second Shannon criterion). n is an integer describing the position of a sample in $s(n\tau)$ and W is defined as:

$$W \equiv \exp(-2i\pi/N) \tag{02}$$

The minus sign in (02) is a convention and must be replaced by + in the reverse Fourier transform. One sees that projection (01) has the mathematical form of a correlation between our signal $s(n\tau)$ and infinite sine functions at frequencies expressed by W. Since there is no phase or time term appearing in (01), the correlation is called zero-lag. Thus (01) can be interpreted as representing the computation of correlations between different sine functions and our signal; $S(k\nu)$ would be the values of these correlations for frequencies $0, \nu, \ldots (N-1)\nu$. We can write (01) under the following matrix form :

$$\begin{bmatrix} S(0) \\ S(\nu) \\ S(2\nu) \\ S(3\nu) \\ \vdots \end{bmatrix} = \tau \begin{bmatrix} 1 & 1 & 1 & 1 & \cdots \\ 1 & W & W^2 & W^3 & \cdots \\ 1 & W^2 & W^4 & W^6 & \cdots \\ 1 & W^3 & W^6 & W^9 & \cdots \\ \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix} \cdot \begin{bmatrix} s(0) \\ s(\tau) \\ s(2\tau) \\ s(3\tau) \\ \vdots \end{bmatrix} \tag{03}$$

(03) can be written $d = \mathbf{G}m$, which is the traditional expression of a direct linear problem (Menke, 1989; Tarantola, 2005), in which **G** is a Vandermonde matrix with ascending diagonal. **G** possesses remarkable symmetry properties imposed by the definition of W (all complex exponentials being equal to each other modulo 2π. Let us, at this point, emphasize that a spectral analysis and a SSA are two very different things, since the two matrices underlying the two methods are completely different.

We next discuss the link between spectral analysis and statistics, or rather probabilities. The convolution between two functions *f(t)* and *g(t)* is expressed by:

$$[f * g](t) \equiv \int_{-\infty}^{+\infty} f(\tau)g(t-\tau)d\tau = \int_{-\infty}^{+\infty} f(t-\tau)g(\tau)d\tau \tag{04}$$

This is also the general formulation of the continuous Fourier transform ($\int_{-\infty}^{+\infty} f(t)\mathrm{e}^{-i\omega t}dt$). Convolution appears in probability theory in the following way. Let α and β be two independent random variables with respective probability densities a(α) and b(β). The probability density g(γ) of the variable $\gamma = \alpha + \beta$ is given by the convolution:

$$g(\gamma) = \int_{-\infty}^{+\infty} a(\gamma - \xi)b(\xi)d\xi \tag{05}$$

This property, associated with the central limit theorem, explains why the normal law occupies a special place in statistics.

We now recall that the auto-covariance (or self-covariance) of a stochastic process allows one to characterize the linear tendencies that exist within this process. Given a signal *x(t)*, the self-covariance at time τ is the mean product:

$$R_x(\tau) = \overline{x(t)x(t+\tau)} \tag{06}$$

In spectral analysis, one hopes to write *x(t)* as a sum of sine functions (see equations (01) and/or (03)), and it can then be written:

$$x(t) = \Sigma_n a_n \cos(\omega_n t + \varphi_n) \tag{07}$$

In that case, (06) becomes:

$$R_x(t) = \Sigma_n \frac{a_n^2}{2} \cos(\omega_n t) \quad (08)$$

(08) contains the same frequencies as the signal *x(t)*. Amplitudes are squared and phases (φ_n) have disappeared. The Fourier transform of (08) is called the spectral density *S(f)*.

This is a very short summary of the links between spectral analysis and statistical or probability theory. If the reviewer wants to know more about the statistics of our analysis, he/she can find the values of amplitudes between parentheses in our Table (since energies are conserved from data space to dual space, whichever decomposition is involved).

We now come to the algorithm of Singular Spectral Analysis. Paradoxically, this method is not a spectral analysis in a strict sense. It has been popularized in the geophysical community by Vautard and Ghil, (1989), Ghil and Vautard, (1991) and Vautard et al. (1992). These authors developed it with the aim of making up for the limitations of the Fourier transform (FT): detecting quasi-periodic components as opposed to perfectly periodic ones in FT, analysis of signals with irregular sampling, or mending gaps in the data. Their work was mainly focused on climatic and paleo-climatic series. The method is curiously under-used in our view. We direct the reader interested in knowing all the details of the method to Golyandina and Zhigljavski (2013), a book with the full, robust mathematical treatment of SSA.

SSA can be summarized in four steps. Let us consider a discrete (non zero) time series ($\mathcal{X}_N$) of length N (N>*2*):

$$\mathcal{X}_N = (x_1, \ldots, x_N) \quad (09)$$

Step 1 (embedding step)

$\mathcal{X}$ is divided into K segments of length L in order to build a matrix $\mathbf{X}$ with dimension $K \times N$ with $K = N - L$+1 will condition our decomposition. This is the first “tuning knob”. Integrating in $\mathbf{X}$ yields a Hankel matrix:

$$\mathbf{X} = \begin{pmatrix} x_1 & x_2 & x_3 \cdots & x_K \\ x_2 & x_3 & x_4 \cdots & x_{K+1} \\ x_3 & x_4 & x_5 \cdots & x_{K+2} \\ \vdots & \vdots & \vdots \ddots & \vdots \\ x_L & x_{L+1} & x_{L+2} \cdots & x_N \end{pmatrix} \quad (10)$$

whereas matrix (03) is a Vandermonde matrix. Comparing the two methods is meaningless.

Embedding, the first step in a SSA, consists in projecting the one-dimensional time series $\mathcal{X}_N$ in a multidimensional space of series *(X_1, ... , X_k)*, such that vectors $X_i = (x_i, \ldots, x_{i+L-1})^t$ belong to $\mathcal{R}^L$, where $K = N - L + 1$. This definition was introduced by Mané (1981) and Takens (1981), with the aim to build a space in which strange attractors could be described correctly, often a Banach space. A strange attractor is an object whose dynamical properties may change and evolve into chaos (that is by nature non-linear). The parameter that controls the embedding is L, the size of the analyzing window, an integer between 2 and $N - 1$. The Hankel matrix has a number of symmetry properties: its transpose $\mathbf{X}^t$, called the trajectory matrix, has dimension K . Embedding is a compulsory step in the analysis of non linear series. It consists formally in the empirical evaluation of all pairs of distances between two offset vectors, delayed (lagged) in order to calculate the correlation dimension of the series. This dimension is rather close to the fractal dimension of strange attractors that could generate that type of series. In this case, it is advised to select for the size of window L very small values (resp. very large K). On the contrary, for SSA, L must be sufficiently large, so that each vector contain an important part of the information contained in the initial time series . From a mathematical point of view, one must work in the frame of *Structural Total Least Squares* for a Hankel matrix (e.g. Lemmerling and Van Huffel, 2001). This is quite different from the fractal dimension mentioned above. Another benefit from considering very large L is the possibility of considering sub-vectors X_i as independent sub-series with distinct dynamics, and thus to be able to identify common characteristics in collections of these sub-series.

Step 2 (decomposition in singular values - SVD)

SVD (Golub et Reinsch, 1971) of non-zero trajectory matrix $\mathbf{X}$ (dimensions $L \times K$) takes the shape:

$$\mathbf{X} = \sum_{i=1}^{d} \sqrt{\lambda_i} U_i V_i^t \tag{11}$$

where the eigenvalues $\lambda_i (i = 1, \ldots, L)$ of matrix $\mathbf{S} = \mathbf{X}\mathbf{X}^T$ are arranged in order of decreasing amplitudes. Eigenvectors U_i and V_i are given by :

$$V_i = \mathbf{X}^T U_i / \sqrt{\lambda_i} \tag{12}$$

The V_i form an orthonormal basis and are arranged in the same order as the λ_i. Let $\mathbf{X}_i$ be a part of matrix $\mathbf{X}$ such that :

$$\mathbf{X}_i = \sqrt{\lambda_i} U_i V_i^t, \tag{13}$$

Embedding matrix $\mathbf{X}$ can then be represented as a simple linear sum of elementary matrices $\mathbf{X}_i$. If all eigenvalues are equal to 1, then decomposition of $\mathbf{X}$ into a sum of unitary matrices is :

$$\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2 + \ldots + \mathbf{X}_d \tag{14}$$

d being the rank of $\mathbf{X}$ ($d = rank\mathbf{X} = max\{i|\lambda_i > 0\}$). SVD allows one to write $\mathbf{X}$ as a sum of d unitary matrices, defined in a univocal way.

Let us now discuss the nature and the characteristics of the embedding matrix. Its rows and columns are sub-series of the original time series (or signal). The eigenvectors U_i and V_i have a time structure, and they can be considered as a representation of temporal data. Let $\mathbf{X}$ be a suite of L lagged parts of $\mathcal{X}$ and $(X_1, \ldots, X_K)$ the linear basis of its eigenvectors. If we let:

$$Z_i = \sum_{i=1}^{d} \sqrt{\lambda_i} V_i \tag{15}$$

with $i = 1, \ldots, d$, then (13) can be written:

$$\mathbf{X} = \sum_{i=1}^{d} U_i Z_i^t \tag{16}$$

that is for the j^{th} elementary matrix:

$$X_j = \sum_{i=1}^{d} z_{ji} U_i \tag{17}$$

where z_{ij} is a component of vector Z_i. This means that vector Z_i is composed of the ith components of vector X_j . In the same way, if we let:

$$Y_i = \sum_{i=1}^{d} \sqrt{\lambda_i} U_i \tag{18}$$

we obtain for the transposed trajectory matrix :

$$X_j^t = \sum_{i=1}^{d} U_i Y_i^t \tag{19}$$

that corresponds to a representation of the K lagged vectors in the orthogonal basis $(V_1, \ldots, V_d)$. One sees why SVD is a very good choice for the analysis of the embedding matrix, since it provides two different geometrical descriptions.

Remark 1: there are very strong common features between performing a SVD of the trajectory matrix, as is done in SSA, and multivariate analyses in principal components (PCA) or Karhunen-

Loève (KL) decompositions, that are commonly found in time series analysis. SSA differs through the nature of its trajectory matrix, a Hankel matrix with a particular structure : its rows and columns are sub-parts of the signal to be analyzed. They both have a meaning, obviously a temporal physical meaning. Such is not the case for PCA or KL.

Remark 2 : In general, the orthonormal basis (U_i) associated to the trajectory matrix and obtained through SVD can be replaced by any orthonormal basis (P_i). In that case, equation (14) becomes $X_i = P_i Q_i^t$ with $Q_i = X^t P_i$, for those who want to perform a Fourier analysis with an orthonormal basis made of sine functions. A classical example of an alternate basis is that of the eigenvectors of a self-covariant matrix (Toeplitz SSA).

Step 3 (reconstruction)

As we have seen, $\mathbf{X}_i$ matrices are unit matrices, and (as in the classical approach) one can "re-group" these matrices into a physically homogeneous quantity (or energetically homogeneous, etc…). This is the second "tuning knob" of SSA. In order to regroup the unit matrices, one divides the set of indices $i\{1, \ldots, d\}$ into m disjoint subsets of indices $\{I_1, \ldots, I_m\}$.

Let *I* be the grouping of *p* indices of $I = \{i_1, i_2, \ldots, i_p\}$; because (14) is linear, then the resulting matrix $\mathbf{X}_I$ that regroups indices *I* can be written:

$$\mathbf{X}_I = \mathbf{X}_{I1} + \mathbf{X}_{I2} + \ldots + \mathbf{X}_{Im} \tag{20}$$

This step is called regrouping the eigen-triplets (λ, U and V). In the limit case $m = d$, (20) becomes exactly (14), and we find again the unit matrices.

Next, how can one associate pairs of eigen-triplets? This means separating the additive components of a time series. One must first consider the concept of separability.

Let $\mathcal{X}$ be the sum of two time series $\mathcal{X}^{(1)}$ and $\mathcal{X}^{(2)}$, such that $x_i = x_i^{(1)} + x_i^{(2)}$ for any $i \in [1, N]$. Let L be the analyzing window (with fixed length), $X, X^{(1)}$and $X^{(2)}$ the embedding matrices of series $\mathcal{X}$, $\mathcal{X}^{(1)}$ and $\mathcal{X}^{(2)}$. These two sub-series are separable (even weakly) in equation (14) if there is a collection of indices $\mathcal{I} \subset \{1, \ldots, d\}$ such that $\mathbf{X}^{(1)} = \sum_{i \in \mathcal{I}} \mathbf{X}_i$, respectively if there is a collection of indices such that $\mathbf{X}^{(1)} = \sum_{i \notin \mathcal{I}} \mathbf{X}_i$.

In the case when separability does exist, the contribution of $\mathbf{X}^{(1)}$ for instance corresponds to the ratio of associated eigenvalues ($\sum_{i \in \mathcal{I}} \lambda_i$) to total eigenvalues ($\sum_{i=1}^{d} \lambda_i$) .

Still in the case of (14), let $\mathcal{I} = \mathcal{I}_1$ be the set of indices corresponding to the first time series, with corresponding matrix $X_{\mathcal{I}_1}$. If both this matrix and that corresponding to the second time series ($\mathbf{X}_{\mathcal{I}_2} = \mathbf{X} - \mathbf{X}_{\mathcal{I}_1}$) are close or identical to a Hankel matrix, then the time series are approximately or perfectly separable. So, regrouping SVD components can be summarized by the decomposition into several elementary matrices, whose structure must be as close as possible to a Hankel matrix of the initial trajectory matrix (this is true on paper only, in reality things are much more difficult).

Step 4 (the diagonal mean, aka hankelization step)

The next, final step consists in going back to data space, that is to calculate time series with length *N* associated with sub-matrices $\mathbf{X}_I$. Let $\mathbf{Y}$ be a matrix with dimension $L * K$ and for each element $y_{i,j}$ we have $1 \leqslant i \leqslant L$ and $1 \leqslant j \leqslant K.$ Let L^* be the minimum and K^* be the maximum. One always has $N = L + K - 1.$ Finally, let $y^*_{ij} = y_{ij}$ if $L < K$ and $y^*_{ij} = y_{ji}$ otherwise. The diagonal average applied to k^th^ index of time series y associated with matrix $\mathbf{Y}$ gives:

$$y_k = \begin{cases} \frac{1}{k} & \sum_{m=1}^{k} y^*_{m,k-m+1} & 1 \leqslant k \leqslant L^* \\ \frac{1}{L^*} & \sum_{m=1}^{L^*} y^*_{m,k-m+1} & L^* \leqslant k \leqslant K^* \\ \frac{1}{N-K+1} & \sum_{m=k-K*+1}^{N-K*+1} y^*_{m,k-m+1} & K^* \leqslant k \leqslant N^* \end{cases} \tag{21}$$

(21) corresponds to the mean of the element on the anti-diagonal $i + j = k + 1$ of the matrix. For $k = 1, y_1 = y_{1,1}$. For $k = 2, y_2 = (y_{1,2} + y_{2,1})/2$, etc…

Thus, one reconstructs the time series with length *N* from the matrices of step 3. If one applies the diagonal mean to unit matrices, then the series one obtains are called elementary series. Note that one can extend naturally the SSA of real signals to complex signals. One only has to replace all transposed marks by complex conjugates.

As we have shown rather quickly, there is in a strict sense no mathematical link between spectral analysis and SSA (despite the names). We acknowledge the fact that SSA should not have been called singular “spectral analysis”.

REFERENCES USED IN NOTE II ABOVE

Claerbout, J. F. (1976), “Fundamentals of geophysical data processing*”,*(Vol. 274), *McGraw-Hill, New York.*

Bracewell, R. N., & Bracewell, R. N. (1986), “The Fourier transform and its applications*”,* (Vol. 31999, pp. 267-272), *McGraw-Hill, New York.*

Golub, G. H. et C. Reinsch (1971), “Singular value decomposition and least squares solutions”, *Linear Algebra*, pp. 134–151. Springer.

Fourier, J. (1822), “Théorie analytique de la chaleur”, par M. Fourier, *Chez Firmin Didot, père et fils*.

Kay, S. M., & Marple, S. L. (1981), “Spectrum analysis - A modern perspective”, *Proceedings of the IEEE*, *69*(11), 1380-1419.

Lemmerling, P. and S. Van Huffel (2001), “Analysis of the structured total least squares problem for Hankel/Toeplitz matrices”, *Numerical Algorithms*, 27 (1), 89–114. 162

Mañé, R. (1981), “On the dimension of the compact invariant sets of certain non-linear maps”, *Dynamical systems and turbulence*, Warwick 1980, pp. 230–242. Springer.

Menke, W. (1989), “Geophysical data analysis: discrete inverse theory”, (Vol. 45), *Academic press.*

Shannon, C. E. (2001). A mathematical theory of communication.*ACM SIGMOBILE mobile computing and communications review*,*5*(1), 3-55.

Prony, R. (1795),”Essai expérimental et analytique: sur les lois de la dilatabilité des fluides élastiques et sur celles de la force expansive de la vapeur d'alcool, à différentes températures”, *Journal de l'école Polytechnique Floréal et Plairial, an III (1795)*, volume 1, cahier 22, 24-76.

Shannon, C. E. (1948), “A mathematical theory of communication”, *The Bell system technical journal*, *27*(3), 379-423.

Stoica, P., and Moses, R. L., 2005, "Spectral analysis of signals", *Prentice Hall edition, New Jersey*.

Takens, F. (1981), “Detecting strange attractors in turbulence”, In *Dynamical systems and turbulence, Warwick 1980* (pp. 366-381). Springer, Berlin, Heidelberg.

Tarantola, A. (2005), “Inverse problem theory and methods for model parameter estimation”, In *Society for Industrial and Applied Mathematics*.

Vautard, R., & Ghil, M. (1989), “Singular spectrum analysis in nonlinear dynamics, with applications to paleoclimatic time series”, *Physica D: Nonlinear Phenomena*,*35*(3), 395-424.

Vautard, R., Yiou, P., & Ghil, M. (1992), “Singular-spectrum analysis: A toolkit for short, noisy chaotic signals”, *Physica D: Nonlinear Phenomena*, *58*(1-4), 95-126.

END OF NOTE II

5: Indeed, the authors seem to be unaware that white noise contains “all frequencies”.

Ironic or somewhat insulting. What is the use of this remark in the review? We do not understand. Either the reviewer considers that the data (which are not ours) are actually white (or pink, or…) noise, or he/she has not seen the Table with the values (between parentheses) of absolute amplitudes of the extracted components that are clearly not identical (hence one cannot speak of white noise).

6: Therefore, finding frequencies in a signal is NOT (has never been) the key point of spectral analysis.

Did we understand this comment? Spectral analysis is not about the determination of the frequency contents of a time series? It may not always be "the key point" but it certainly is "a key point". "Spectral analysis" is often called "frequency analysis".

7: The only interesting question is to estimate their statistical significance.

The "only" interesting question? When projecting a signal on an orthogonal basis, such as in Fourier spectral analysis, the statistics are certainly not the main goal.

8: But there is simply no mention at all of statistics in this paper.

Indeed, in this paper we do not need a statistical analysis, except for the rough estimate of uncertainties of the central frequency of individual SSA components. If there were large numbers of spectral peaks/SSA components and only some could be attributed to planetary frequencies, we could not decide which ones were significant. But when most observed frequencies are recognized as naturally present in the series that is analyzed, it is unreasonable to assign it to chance.

As we have recalled above in Note II, classical spectral analysis and SSA are not statistical analyses, even if one can link these two different branches of mathematics. If the reviewer is asking about the meaning of periodicities we identify, then he/she must have the same concern about the work of Hays et al (1976) or Vautard and Ghil (1989). We actually share part of their lists of periods; not only do we confirm but we extend their lists.

9: SA is now a very classical spectral analysis method. It was largely developed in the 1980s (Broomhead and King, 1986) and was used for paleoclimatic studies since then (Vautard and Ghil, 1989).

We know and cite this useful paper by V and G (the reviewer?). Indeed SSA has been used in climate studies but not in a number of fields and series we are now analyzing. If the method is so classical, why the questions on the statistics of our results? Could the reviewer help us by giving us the references to papers that perform statistical tests on the components of a SSA analysis? In any case, we have discussed this at length in Note II above. Now, if the question is about the small absolute percentages we indicate in our Table, asking about the reality of the components we extract, there are two answers. The first is that this percentage is the fraction of the total signal. But the largest (first) component is in general the trend, which is non linear. It would be legitimate to de-trend the series, since we are interested in periodic and quasi-periodic components, and the amplitudes of all the other components would be correspondingly increased. The second answer is that indeed several of these components are very small compared to the general trend: to which extent could FA as well as SSA (i.e. SVD) be able to separate information in these signals? This is not anymore a question of algorithm (stricto sensu), but of numerical resolving power and precision of computation. This means of course computer, operating system and language (eg. *https://www.boost.org/doc/libs/1_47_0/libs/math/doc/sf_and_dist/html/math_toolkit/special/sf_poly/sph_harm.html*). We have used Matlab on two Xeon 7 processors, thus MKL and 64 bits under Ubuntu (Linux). We hope this answers the reviewer's concerns about the precision of our results (which is on the order of 10^{-16}) and the fact that our results are exact, since they rely on mathematical functions (SVD, Hankel, FFT, etc …) implemented by the 5000 mathematicians, numericists and computer scientists that work in the company (Mathworks) that sells Matlab.

10: It was even applied to the Vostok record as early as 1994 (Yiou et al., 1994).

True and we do not contradict the results, we complement them.

11: Again, it is strange not to find these references in a scientific paper on SSA applied to paleoclimate records in general and to the Vostok data in particular.

Yiou et al. (1994) find with SSA the same components found by Petit and Jouzel (1977) using Fourier analysis. The papers mentioned by the reviewer, e.g. Vautard and Ghil (1989), do not describe the full mathematical apparatus of SSA, but are a very good introduction to the method. We can make a comparison with the discovery of wavelets by Jean Morlet (Morlet et al., 1982). A rudimentary tool first developed by a geophysicist, then developed much more fully by mathematicians who were specialists in time series analysis (eg. Grossmann and Morlet, 1984), up to the pyramidal algorithm of Mallat (2009) that is now (as is SVD) in the top ten of the most used algorithms in research (*cf.* Cipra, 2000). In the same way, the work of Vautard and Ghil barely touches on the concept of separability of information, as we defined it in Note II above. This is the reason why we systematically cite Golyandina and Zhigljavsky's (2013) book. We have detected more periods because, as stated in our paper, we have used a variant of SSA, called iterative SSA (or iSSA). This implies changing the size of the analysing window at each step of detection and extraction of a period (component) so that equation (14) remains as close as possible to a Hankel matrix. Indeed, there are a large number of variants of SSA: iterative SSA (e.g Golyandina and Zhigljavsky, 2013), varimax Rotation SSA (e.g. Portes et Aguirre, 2016), multitaper SSA (Dettinger et al., 1995), circulating SSA (e.g. Bogalo et al., 2021), etc … All aim to ensure optimal separability as a function of the information carried by a given signal (time series). Vautard and Ghil (1992) had performed the simplest, most basic kind of SSA, so the differences in results (improvements) are not surprising.

REFERENCES FOR point 11

Bógalo, J., Poncela, P., & Senra, E. (2021). Circulant Singular Spectrum Analysis: A new automated procedure for signal extraction. *Signal Processing*, *179*, 107824.

Cipra, B. A. (2000). The best of the 20th century : Editors name top 10 algorithms. SIAM news 33 (4), 1–2. 73

M.D. Dettinger, M. Ghil, C.M. Strong, W. Weibel, P. Yiou, Software expedites singular-spectrum analysis of noisy time series EOS, Trans. Am. Geophys. Union, 76 (1995), pp. 12-21

Golyandina, N., and Zhigljavsky, A.,: Singular Spectrum Analysis for time series, Vol. 120, Berlin, Springer, 2013

Grossmann, A. and J. Morlet (1984). Decomposition of hardy functions into square integrable wavelets of constant shape. SIAM journal on mathematical analysis 15 (4), 723–736. 134

Mallat, S. G. (2009). A theory for multiresolution signal decomposition: the wavelet representation. In *Fundamental Papers in Wavelet Theory* (pp. 494-513). Princeton University Press.

Morlet, J., G. Arens, E. Fourgeau, et D. Glard (1982). Wave propagation and sampling theory —part i : Complex signal and scattering in multilayered media. Geophysics 47 (2), 203–221. 134

Portes, L. L., & Aguirre, L. A. (2016). Matrix formulation and singular-value decomposition algorithm for structured varimax rotation in multivariate singular spectrum analysis. *Physical Review E*, *93*(5), 052216.

12: In fact, SSA has been applied to Vostok and to LR04 in many previous papers over the last 20 years : it is so classical that this analysis is now even part of some student textbooks (eg. Martinson 2018, page 531, Figure 15.6 showing the SSA analysis of LR04). In any case, it is difficult to find any novelty in this paper.

The many papers that have all found the same set of 4 or 5 periodicities need not be all cited! The point of our paper is that we find more components and can identify them. Also, this is a way of testing our results on well-known series before applying them to new complementary and/or longer, better sampled series. We find more “Milankovitch”-type periods.

13: Spectral analysis is a branch of statistics.

No! It can use tools from the field of statistics, it is not an (ancillary) branch of this field. One can use a number of statistical tools with Fourier, not with wavelets or SSA, or PCA. One can make statistical inferences with Fourier (autocorrelation, DSP). SSA is a decomposition on an ad-hoc basis, not sinus functions (and therefore not Laplacian). Same for wavelets. See Note II.

14: Finding many frequencies (real or spurious) in a signal is known to be extremely easy, since the discovery of the Fourier transform. In particular, finding astronomical frequencies and their numerous harmonics in paleoclimatic signals is certainly not surprising.

Yes but whereas Fourier creates harmonics (Sinc function), SSA a priori cannot: it is difficult to create a harmonic of a signal which is not periodical stricto sensu. See Note II.

15: Actually, the opposite would be much more unexpected. This is even more true for the insolation itself !…

Why for insolation, that is calculated from solar irradiance (a constant), eccentricity, obliquity and precession that are all cyclical?

16: Not finding the main and secondary astronomical periodicities that are used as input of the insolation computation would be a sure signature of some methodological error. I therefore do not understand the value of applying a spectral analysis method (in particular on insolation), without discussing the "detection level" problem.

The secondary astronomical periodicities are not input but actual observations! We fail to see the link with the previous comment. As already stated above, the detection level is a concept that is computer-dependant (32 vs 64 bit, Linux vs Windows, MKL vs Matlab, etc …) and not methoddependant. An example is given in the link below, that shows how the actual realization of a decomposition depends on the computer, the library, the software and the nature of the processor.

https://www.boost.org/doc/libs/1_47_0/libs/math/doc/sf_and_dist/html/math_toolkit/special/sf_poly/sph_harm.html

17: Indeed, the key mathematical difficulty is to decide whether it is “noise” or “signal” using some statistical test. There is no mention of any statistics in this paper, which certainly disqualifies it entirely. This is strange since standard SSA tools used in the climate community (eg. SSA-MTM toolkit) do provide some statistical tools and safeguards.

Again on statistics. See above and Note II. As soon as one exits from the (restricted) climate community, people use it in the same way we do (see references above). Statistics is not what qualifies a signal as being noise or not. By definition, a noise is a signal (stochastic or not) that is superimposed on a signal of interest. For instance, when one performs electrical measurements in urban zones, one often adds a 50 Hz signal, that is an almost perfect sine function. Which statistics will tell us that the 50 Hz signal is a noise?

18: The harmonics and combination tones found by the authors have often a very small amplitude (below 1% or 0.1% of the variance) and to prove their significance is certainly not an easy task, given that the records are rather “short” and “noisy”. This is precisely the main difficulty of spectral analysis, and a problem that climate scientists are usually facing. But the authors of this paper are not even mentioning or discussing this central point, as if they were unaware of the main problem of “spectral analysis estimation”.

These small amplitudes must be discussed in parallel with the trend (see point 9). Which as we argue could be a “short” segment of a component with a quasi-period much longer than the interval of time over which the time series is known. And if we let aside the trend, then most of the components (quasi-periods) fit the characteristic astronomical quasi-periods. This analysis is not necessarily “difficult”; see the critical sampling intervals in Papoulis (1977) or Claerbout (1984). Moreover, we have already answered this point (Note II).

19: In addition, it is worth mentioning that detecting several significant frequencies (let alone NUMEROUS significant frequencies) is a much more difficult problem than detecting a single one.

Point already made, already answered. This might have been problematic in the ‘60s, but not any more. One can now reconstruct a hushed voice with a mobile phone (using piezo electricity thus frequencies).

20: Indeed, the null-hypothesis cannot be a simple (red or white) noise hypothesis.

What is the point?

21: A classical work-around is to perform Monte-Carlo simulations of synthetic signals with and without each single periodicity and assess whether the results are statistically distinguishable. But this would probably be a very different paper.

If there are synthetic signals, then there is a model behind. In the present paper, we do not introduce any model but we perform a signal analysis of cores that have already been analyzed in several papers and in which NO model is necessary.

We do not use Monte-Carlo since the work by Kirkpatrick (1977) on Simulated annealing. It is much more efficient (and elegant) than a simple distribution of models. In signal analysis, statistics underlie the matrices and decompositions we use (e.g. Mañé, 1981 ; Takens, 1981 ; or Lemmerling and Van Huffel, 2001). If the reviewer wishes more information, in order to give SVD (that we used in SSA) a statistical meaning, we refer her/him to Muller et al. (2004) and see Note II.

REFERENCE TO POINT 21

Muller, N., Magaia, L., & Herbst, B. M. (2004). Singular value decomposition, eigenfaces, and 3D reconstructions. *SIAM review*, *46*(3), 518-545.

22: Just to be more precise: Of course, I do not mention the uncertainty in the periodicities as listed in table 1 (obtained after SSA decomposition and resampling of the data), but the significance of the SSA components themselves. In particular, the KEY parameter here is the choice of the embedding dimension. If chosen too high (extremely likely here), then there are a lot of SSA components with dubious relevance. The same phenomenon applies to all spectral methods: the more degrees of freedom you are allowing, the more frequencies you get, but the less significant they are. What is the embedding dimension of the analysis here? This is not even mentioned ! But according to line 62, “SSA was able to extract up to 75 cycles”. My guess is that the embedding dimension is a few hundreds, which represents quite a lot of freedom for the method to generate a lot of “spurious” frequencies.

We agree with the first part of this comment. But in our case the embedding dimensions (given the size of the time windows over which the series are defined) are small! In iterative SSA (iSSA) they change at each step. Should we add a table with the window sizes for each component successively extracted? The total of 75 is for the 8 series analyzed. For each one (looking at the columns of the Table) the number is far smaller: there are not 75 lines per column!

23: It is very strange to use the old Vostok record (Petit et al. 1999) which covers only 4 glacial cycles while a more recent one (Dome C), covering twice as many cycles, has been available for more than a decade now (see eg: Guo et al (2012) for the SSA analysis of CH4 at Dome C over 8 cycles, not 4 as here).

It is not strange: it is one of the most quoted sequences. We wanted to check our method and thought the reader might be interested by this comparison. Of course (if decided by the editor), we could do the analysis for the Guo et al (2012) series. But in our view, it is not needed in the frame of the present paper.

24: Obviously, the results are likely to be statistically more robust with a longer time series, if statistics matters of course. “Finding not only the main expected Milankovitch periodicities but also many “secondary” components with much smaller amplitudes gives confidence in our iterative SSA method”. But I would expect standard methods to do the same, with attributing to these periodicities a very low confidence. In any case again, finding many frequencies is easy and NOT the main problem of “spectral analysis”. Besides, what is the subject of the paper? Is it about climate or mathematics?

What does this mean? Why separate both? We use mathematical techniques to test and check their power in analyzing actual observational time series commonly associated with climate and its changes. Otherwise, see Note II.

25: The authors appear unaware that the word “insolation” is not sufficient to describe the astronomical forcing of climate. They are presenting the spectral analysis of “the insolation of Laskar et al (2004)”. This is clearly a very insufficient description. When looking more closely at the spectrum (Figure 3), it seems that they are in fact discussing global averaged insolation (linked to eccentricity only)! Otherwise the spectrum would be dominated by the 41k and the 23k periodicities. Since there is no power in the obliquity and precession band, the only explanation is that the authors are using global averaged insolation. But this is a major misunderstanding on the Milankovitch forcing. Indeed, it is well known since the 19th century that global insolation changes are much too weak to affect climate and scientists have discussed local seasonal insolation, not mean global one. When climate scientists talk about "insolation" in the context of glacial cycles, they are talking about summer insolation at high northern latitude according to Milankovitch theory. THIS IS NOT what is used in this analysis.

See Note II.

26: In other words, this paper has missed entirely its claimed goal of discussing the Milankovitch forcing, simply by looking at the wrong insolation…

To conclude, its seems to me that the authors have misunderstood both the aim of "spectral analysis" and the foundations of the "Milankovitch theory". I cannot be favorable to publication.

On these two points, we have shown that the reviewer is wrong. We have specifically devoted the two longer Notes I and II above to explain where and how. We appeal to the Editor and confirm our submission.