Article

Northern Hemisphere Glaciation: Its Tectonic Origin in the Neogene Uplift

Hsien-Wang Ou

Lamont-Doherty Earth Observatory, Columbia University, Palisades, NY 10964, USA; hwo1@columbia.edu

Abstract: The Earth has cooled since the early Pliocene, which was punctuated by accelerated cooling indicative of thresholds. I posit that the cooling was initiated when the Neogene uplift of the Tibetan highland caused it to ice over, augmenting the albedo. I formulate a minimal warm/cold/Arctic box model to test this hypothesis and prognose the Pliocene climate. In particular, based on model physics, I discern three thermal thresholds as Pliocene cools: (1) when the Arctic temperature falls below the marking temperature of the ice front, the East Greenland ice sheet would descend to the sea level and calve into the Nordic Seas; (2) when the Arctic temperature cools to the freezing point, the ice sheet would form and expand over circum-Arctic lowlands to cause a massive deposition of ice-rafted debris marking Northern Hemisphere glaciation (NHG); (3) when glacial state persists through low eccentricity, it would cause a transition from obliquity- to eccentricity-dominated glacial cycles. Aligning these thresholds with the observed ones around 3.5, 2.7, and 1 million years ago, the model produces a temporal evolution of the Pliocene temperature as well as its driving albedo change. Since the latter can be accommodated by the observed one, it supports the Neogene uplift as the tectonic origin of NHG.

Keywords: Northern Hemisphere glaciation; Greenland ice sheet; Neogene uplift; Pliocene cooling; glacial cycles

1. Introduction

The latest cooling of the Earth on the tectonic timescale began in the early Pliocene about 4 million years ago (Ma) [1,2]. It was punctuated by glaciation of West Greenland [3,4] before culminating in Northern Hemisphere glaciation (NHG) manifested in a massive increase of ice-rafted debris (IRD) [5]. Possible origins of the Pliocene cooling have been widely explored and extensive critiques of some proposed mechanisms have been put forth. Rather than reiterating these criticisms, I shall raise below additional points that are often overlooked.

First, one should be cautious in assigning causative roles to internal properties of the climate system in probing its response. One prominent example is the partial pressure of the atmospheric carbon dioxide ($pCO_2$), which would equilibrate with temperature over millennia [6]; hence, it may not be independently prescribed. As a case in point, the two track each other closely during glacial cycles [7], but since the temperature signal is induced by the orbital forcing via energy balance—with or without $pCO_2$—the latter signal is a response rather than a driver of the temperature change. Over the tectonic timescale, $pCO_2$ has decreased to the preindustrial level since 20 Ma [8] and stayed low through the Pliocene, so it cannot account for the observed cooling of about $10 \degree C$ [4] based on the calculated sensitivity [9].

One well-known tectonic event is the closure of the Panama Isthmus, which has warmed the subtropics and increased poleward moisture transport [10], with the latter argued to propel the ice growth despite the warmth [11]. However, this causality is opposite to that found in observations and climate models [12,13] hence remains unsupported. Moreover, the seaway closure was largely complete around 3.6 Ma, which preceded NHG...
by 1 Myr [14], and, given the instantaneous (decadal) ocean response, this timing mismatch is difficult to bridge. As such, the seaway closure may cause only a temporary disruption of the Pliocene cooling [15].

Another tectonic candidate is the widespread uplift of the northern continents during the Neogene [16], which may enhance weathering to draw down $pCO_2$, thus cooling the planet [17], but the effect is likely minor [18] and, in any case, weathering constitutes negative feedback to further muddle causality. An additional effect of the uplift well demonstrated by climate models is the amplification of the planetary wave [19], which would cool the northern lands, but this is predominately a winter phenomenon [20], while it is the summer temperature that regulates the ice extent [21,22].

Another readily discerned effect of the uplift is the increase in planetary albedo caused by ice-topped highlands. The author of [23] shows that the Tibetan highland was ice-covered during the last ice age, accompanied by about 10 °C cooling, underscoring the efficacy of its albedo effect because of its vast size. This effect, however, is largely overlooked in the literature, and not even mentioned in most reviews of the subject [24,25]. It is to quantify this albedo effect due to uplift and assess its possible role in the Pliocene cooling that prompts the present inquiry.

For this purpose, I shall present a minimal warm/cold/Arctic box model, which is nonetheless thermodynamically closed to allow a prognosis of the Pliocene temperature and glacial regimes. Specifically, because of inherently turbulent planetary fluids, a climate state is a macroscopic manifestation of a nonequilibrium thermodynamic (NT) system, and hence subject to maximum entropy production (MEP), a generalized second law. Readers are referred to [26,27] for reviews of MEP, whose physical basis is continually strengthened by its linkage to the fluctuation theorem [28,29], as the latter has been tested in the laboratory [30]. In addition, recent eddy-resolving and direct numerical simulations of horizontal convection [31,32] have reproduced its signature feature (a mid-latitude front [33]), which is absent from coarse-grained general circulation models [34], and its application has replicated broad climate features and resolved long-standing glacial puzzles [35,36]. With these supports, I am justified to posit MEP as a working hypothesis in formulating our model closure (see further discussion in Section 6).

The organization of the paper is as follows: I first formulate the box model in Section 2 and then link temperature to the radiative forcing in Section 3. In Section 4, I discern the thermal thresholds of changing glacial regimes. Aligning these thresholds with observed ones, I calculate in Section 5 the temporal evolution of the Pliocene climate, which supports the uplift hypothesis. I provide additional discussions in Section 6 and conclude the paper in Section 7.

2. Box Model

I consider a box configuration (Figure 1; all symbols are listed in Appendix A), in which the coupled ocean/atmosphere are divided into warm/cold boxes separated by the subtropical front (30° N) and an Arctic Ocean box appended at the Fram Strait (80° N). The sea-surface temperature (SST; thick solid line) represents a first-order description of the observed one, which may also be justified as an asymptotic state of infinite eddy mixing when subject to differential solar forcing. Such mixing would homogenize the temperature gradient within each box, which would be zero (hence isothermal) in warm/Arctic boxes because of the vanishing gradient at the equator/pole of a sphere, but retain a finite value in the cold box due to its anchoring at a possibly different Arctic temperature, a profile that is nonetheless fully specified by the derived cold-box average (thick dashed line). Since there is no blockage of the Arctic air, the troposphere comprises only two isothermal boxes.

Being a minimalistic model, I retain only dominant energy balances: the system is heated differentially by the incoming shortwave (SW) flux ($q_f^i$), which after atmospheric absorption and reflection by the planetary albedo is absorbed by the ocean ($q_i$, referred as ‘radiative forcing’) to differentiate the SST. The latter, in turn, differentiates the surface-air temperature (SAT) by the convective flux ($q_c$) neglecting by comparison latitudinal
variation of the surface- and outgoing-longwave (LW) fluxes. Because of the thermal inertia of the ocean, I neglect its seasonality. Air, however, is strongly warmed in summer by atmospheric absorption, which remains bounded by the underlying SST in maintaining radiative-convective equilibrium of the troposphere. This ‘convective bound’ is a distinct feature in observation [37] (their Figure 10.7, lower panel), so I shall take SST as a proxy of the co-zonal summer SAT that controls the ablation, and hence glacial margin.

In our convention, global means are overbarred, deviations (referred to as ‘differential’) are primed and subscripted 1/2/3 for warm/cold/Arctic boxes. The latitudinal distance is defined as \( y = \sin(\text{latitude}) \), so equal intervals span equal surface area on a sphere, and the ocean boundaries at 30 and 80° N have \( y = 0.5 \) and 0.98, respectively. Given the smallness of the Arctic box (\( \Delta y = 0.02 \)), I shall neglect its effect on the other box-averaged temperature, which thus is simply the two-box solution derived in [29]. Readers are referred to [29] for mathematical details, but the abbreviated derivation given below should aid the self-containment of the paper.

3. Differential Temperature

In the following, I shall first link the differential SAT to SST via atmospheric MEP. With the dominant heat balance stated earlier, the irreversible entropy production of the atmosphere equals the entropy flux through its lower boundary, which is of the form \( \sigma(T - T_a) / T_a \), where \( \sigma \) is the air-sea exchange coefficient set to 10 Wm\(^{-2}\)K\(^{-1} \) [29]. Given the small temperature deviation from its global mean (in Kelvins), the inverse temperature in the above expression can be linearized, so the differential entropy production (that is, removing the global mean) varies as (the symbol “∼”)

\[
\sigma'_a \sim \sigma(T_2' - T_{a,2}') T_{a,2}',
\]

which is expressed in the cold-box variables as the warm box has the same value. Maximizing this entropy production against \( T_{a,2}' \) yields

\[
T_{a,2}' = T_2' / 2,
\]

a well-known result of MEP [26] because of mutual compensation of the thermodynamic flux (heat transport, the bracketed term) and thermodynamic force (differential temperature).
Having linked SAT to SST, I next apply the oceanic MEP to link SST to the radiative forcing $q'_2$. Since the differential heating of the cold ocean box is the radiative forcing minus the convective flux, the oceanic counterpart to (1) is

$$
\sigma'_o \sim [q'_2 - \alpha (T'_2 - T'_{a,2})] T'_2.
$$

(3)

Applying (2) and maximizing (3) against $T'_2$, we obtain

$$
T'_2 = q'_2 / \alpha;
$$

(4)

the differential SST is thus linear in the differential forcing. As a cursory check, a forcing range of 200 Wm$^{-2}$ [37] would produce an SST range of 20 °C, not unlike that observed [38]. In addition, the model-deduced meridional overturning circulation depends only on $\alpha$ to yield a transport of 11 Sv for the 6000 km wide North Atlantic [36], which is of the same order as the observed one [39]. These observational comparisons constitute a veritable test of MEP as they contain no free parameters—unlike previous comparisons involving tunable parameters [40].

Given the narrowness of the Fram Strait, which is the sole gateway to the Arctic Ocean since the Miocene [41], I shall neglect the heat transport into the Arctic Ocean, so that

$$
q'_2 - \alpha (T'_3 - T'_{a,2}) = 0.
$$

(5)

Substituting from (2) and (4), we derive

$$
T'_3 = \frac{3}{2\alpha} q'_2.
$$

(6)

As seen in Figure 1, the Arctic Ocean could be 50% colder than the averaged subpolar temperature (thick dashed line) with respect to the global mean, not unlike the present case [37] (their Figure 8.2a), and since the Arctic temperature anchors the cold-box SST (thick solid line), the latter has a finite gradient in halting the glacial advance.

Having linked temperature to radiative forcing, I shall next examine how uplift-induced albedo change affects radiative forcing, and hence temperature. The radiative forcing of the cold box is

$$
q_2 = q'_2 (1 - b - a_2),
$$

(7)

where $q'_2$ is the incoming SW flux, $b$ the (constant) atmospheric absorption, and $a_2$ the planetary albedo; only the last is altered by uplift. Let $i$ be the area rendered ice-covered by uplift (in fraction of the cold-box area), then the cold-box forcing varies as

$$
\delta q_2 = -q'_2 \delta a_2 = -q'_2 \Delta r \delta i,
$$

where $\Delta r$, set to 0.7, is the excess albedo of ice over land. Since the Neogene uplift is dominated by mid-latitude lands within the northern cold box, they are assumed to be free of low clouds concentrated in high latitudes [42] (high clouds have little albedo), and with the northern cold box comprising one quarter of the global surface, the global forcing thus varies as

$$
\delta \bar{q} = \delta q_2 / 4,
$$

(9)

and the differential forcing varies as

$$
\delta q'_2 = \delta (q_2 - \bar{q}) = 3 \delta \bar{q}.
$$

(10)

That is, because of the concentration of uplift in the northern cold box, its effect on the differential forcing is three times that on the global forcing. The latter induces global cooling via

$$
\delta \bar{T} = \pi \delta \bar{q},
$$

(11)
where the ‘global sensitivity’ $\bar{s}$ has a range of [0.5, 1.5] K(Wm$^{-2}$)$^{-1}$ based on climate models [9]. Substituting from (9) and (8), the global cooling is thus linked to the ice cover via

$$
\delta T = - (\bar{s} q^2) \delta i / 4,
$$

and from (4), (6), (10), and (11), the global cooling would cool the cold and Arctic boxes via

$$
\delta T_2 = \delta (T + T_2^o) = \left(1 + \frac{3}{\alpha^2}\right) \delta T,
$$

$$
\delta T_3 = \delta (T + T_3^o) = \left(1 + \frac{9}{2\alpha^2}\right) \delta T.
$$

It is seen that the cold box is cooled more strongly than the globe temperature due to contribution from the differential forcing (10), and the cooling is further amplified in the Arctic Ocean because of blockage of the incoming ocean heat.

With the variations derived above, one may calculate the temperature evolution given their initial values and observed uplift, a prognostic scheme that is, however, thwarted by the uncertain uplift history. On the other hand, there is a proxy time series of the subpolar SST, which exhibits abrupt cooling episodes indicative of thresholds, so if we can identify these thresholds from our model, we may then fix their timing to diagnose the uplift history. Taking this approach, I shall next consider the thermal thresholds from our model.

4. Thermal Thresholds

To illustrate the thermal thresholds as the Pliocene cools, I draw in Figure 2 three cold/Arctic-box SST profiles (same as the summer SAT) representing the initial condition of the early Pliocene (EP), the expansion of the Greenland ice sheet (GIS), and NHG. For a constant lapse rate set to $\gamma = 6$ °C, the temperature profiles also represent the snowline height, both are tick-marked on the right ordinate. The highland above the snowline would be ice-covered, and the net accumulation would propel the ice down the slope until the ice flux is depleted by ablation, an ice edge defining the ‘glacial line’ and its temperature $T_g$ referred to as the ‘glacial marking temperature’ (GMT). This temperature is driven in Appendix B, which turns out to be primarily an intrinsic property of the ice sheet, and its deduced value of about 5 °C is consistent with the present observation [43], which is indicated in a dashed line. It is recognized that the GMT derived in Appendix B over a slope would be altered for a sea-level ice sheet [36], but since we are not prognosing the ice extent on the sea level, it is drawn as level for illustration.

Dark-shaded columns represent Greenland and the Tibetan plateau, which were uplifted during the Neogene. Greenland is perhaps elevated in two steps [44], and being located in high latitudes with lower snowline, its eastern part was already ice-covered by the late Miocene [45], and hence had no impact on the subsequent albedo change. For the Tibetan highland, I subscribe to its uplift history synthesized by [46] based on vegetation data, which shows a notable lag of its northern from southern provinces (see also [47]), but despite the timing spread, they converge and reach beyond the snowline in the early Pliocene (EP) to initiate the cooling. It is noted that [48] proposed an earlier uplift (~8 Ma) based on assumed crustal/mantle thickness, which the author acknowledged as speculative, and it conflicts moreover with the early Pliocene uplift (~4.5 Ma) based on sedimentary evidence [49]. Also of relevance is the fact that the Tibetan plateau was ice-covered during the last ice age when the snowline fell below the plateau height to cause about 10 °C cooling, underscoring the potency of its albedo effect because of its vast spatial coverage [23]. Given the phased uplift of the Tibetan highland, one expects a gradual cooling on the tectonic timescale and, as the cooling proceeds, the lowering of the snowline would expose lesser highlands, such as the American West [24] and margins of the North Atlantic Ocean [50], to glaciation [51], so it is this combined uplift area that determines its albedo effect and whether it is sufficient to cause NHG.
Without such seeding, there can be no sea-level ice sheet so long as the surface temperature weakening radiative forcing by cooling the surface, leading to the formation of perennial point or approximately $T_0$ falls below the glacial marking temperature ($T_g$, dashed) to fully glaciate Greenland. The second threshold (labeled NHG for Northern Hemisphere glaciation) is when the Arctic Ocean is cooled to the freezing point, so the polar cap (light-shaded) would form on the circum-Arctic lowlands and advance toward $T_g$.

The first threshold (labeled GIS for Greenland ice sheet) is when the Arctic temperature cools to the GMT (dashed line), so it is defined by

$$GIS : T_3 = T_g.$$  \hspace{1cm} (15)

When this threshold is crossed, the ice sheet over East Greenland would descend to the sea level and advance southward toward GMT, which, however, is cut short by the shoreline to calve into the Nordic Seas, resulting in minor IRD events. It is important to note that this threshold leading to full glaciation of Greenland is predicated on the existing ice sheet atop East Greenland, which provides the seeding for its expansion to sea level. Without such seeding, there can be no sea-level ice sheet so long as the surface temperature remains above the freezing point, leading to the second threshold delineated next.

The second threshold (labeled NHG) is when the Arctic Ocean is cooled to the freezing point or approximately

$$NHG : T_3 \approx 0.$$  \hspace{1cm} (16)

At this juncture, the convective flux has exhausted its capacity in countering the weakening radiative forcing by cooling the surface, leading to the formation of perennial sea ice, which would expand to the whole Arctic Ocean and expel the snowline to the surrounding lowlands. Subsequently, even a slight cooling perturbation would allow a foothold of positive ice mass balance, which would then expand the polar cap southward until its leading temperature reaches GMT (that is, the intersect of solid and dashed lines), as indicated by the light shade. The polar cap would calve into the surrounding ocean to produce massive deposition of IRD—the defining marker of NHG.

The generation of sea/land ice during NHG would augment the albedo to continue the Pliocene cooling, but also significantly, with the polar cap now encroaching on the contiguous northern lands, it would activate the ice-albedo feedback to implement the precession forcing, which otherwise would be nil on account of Kepler’s second law [22]: a lower summer insolation, for example, would be countered by a longer summer to yield zero annual mean in the radiative forcing. As posited by this author [36], this is

![Figure 2. Profiles of cold/Arctic-box summer SAT (doubling as the snowline, both tick-marked on the ordinate) to illustrate thermal thresholds as the Pliocene cools. The initial condition (labeled EP for the early Pliocene) is when the Tibet plateau rises above the snowline to initiate the Pliocene cooling. The first threshold (labeled GIS for Greenland ice sheet) is when the Arctic temperature falls below the glacial marking temperature ($T_g$, dashed) to fully glaciate Greenland. The second threshold (labeled NHG for Northern Hemisphere glaciation) is when the Arctic Ocean is cooled to the freezing point, so the polar cap (light-shaded) would form on the circum-Arctic lowlands and advance toward $T_g$.](image-url)
the reason that the 41 ky obliquity signal prevails prior to NHG, but the 21 ky precession signal would emerge during NHG to be followed by distinctive mid-Pleistocene transition (MPT) to 100 ky ice-age cycles paced by eccentricity [52]. Ref. [36] noted that MPT remains unresolved for lack of evidence of putative pCO2 decrease and regolith removal [52–54] and provided an alternative mechanism of a bistable ocean, to which readers are referred for details. A brief derivation is nonetheless provided below to further quantify MPT and address causality between NHG and glacial cycles [55].

Crudely, ref. [36] postulated that the augmented differential forcing beyond the convective bound (Section 2) would vault the cold box into the freezing point by MEP, rendering a bistable temperate/freezing-point cold box, which would translate to a bistable polar-cap/Laurentide ice sheet (LIS). The bistable forcing interval is set by the global convective flux $\bar{q}_c$, which has a centerline lying about 1.65 $\bar{q}_c$ below the global forcing. The global surface heat balance is of the form

$$q_c = q - q_{LW}, \quad (17)$$

where $\bar{q}$ and $\bar{q}_{LW}$ are the global radiative forcing and the surface LW flux, respectively. Defining a ‘greenhouse’ parameter

$$s_g \equiv \delta q_{LW} / \delta T, \quad (18)$$

and setting it to $s_g \approx -2$ [42] (their Figure 1 over temperature range [283 K, 293 K] in accordance with Figure 3), we derive

$$\delta q_e = \delta q - \delta q_{LW} = \left( 1 - s_g \right) \delta T. \quad (19)$$

The global convective flux thus decreases during the Pliocene cooling due to both the weaker radiative forcing and drier air. As shown in [36], the threshold of MPT is when the ‘excess’ convective flux defined below vanishes, so for the third threshold,

$$\text{MPT: } q_e \equiv 1.65 \bar{q}_c + q'_2 = 0. \quad (20)$$

When this threshold is crossed, the glacial state would persist through low eccentricity, thus allowing ice growth to LIS that defines the ice age, and the glacial cycles would transition from obliquity- to eccentricity-dominated periodicity. Substituting from (20), (19), (10), and (11), the excess convective flux varies as follows:

$$\delta q_e = \left[ 1.65 \left( 1 - s_g \right) + \frac{3}{5} \right] \delta T. \quad (21)$$

Therefore, the MPT threshold is specified by its initial value and global cooling. Having identified these thresholds, we shall next discern them from the observed temperature evolution.

5. Pliocene Climate

Ref. [4] has provided a high-resolution subpolar SST from 4 Ma to the present. The thick line shown in Figure 3 represents my interpretation of its broad features, which can be characterized as gradual cooling punctuated by three more abrupt episodes of cooling, around 3.5, 2.7, and 1 Ma. Although [56] has inferred a mid-Miocene warmth (~3 Ma) from their data spanning 3–1.7 Ma, it is better interpreted, based on the longer time series of [4], as slowing of the cooling when sandwiched between two rapid cooling episodes. Since the temperature is inferred from proxy data of the eastern North Atlantic with summer bias, it is expected to be quite warmer (about 5 °C) than the averaged cold-box temperature ($T_2$) [57], but the abrupt changes dovetail the modeled thresholds to allow fixation of their timing (dashed vertical lines).
Figure 3. Time evolution of the modeled global (\(T\)), averaged cold-box (\(T_2\)), and Arctic (\(T_3\)) temperatures and uplift-induced ice cover (\(i\)) when thermal thresholds (dashed vertical lines) are aligned with abrupt cooling of the observed subpolar SST (thick solid line, adapted from [4]). The shaded portions are solution spans for global sensitivity ranging from 0.5 (thin solid line) to 1.5 (thin dashed line). The solid square is the estimated ice cover by uplift, which is seen sufficient to cause NHG.

For the initial condition, it is noted that the early Pliocene (EP) is slightly warmer than the Holocene [58], so I set the following values based on the present observation [37] (their Figures 6.3, 6.13 and 6.14). Specifically, I set incoming SW fluxes to \(a_1, a_2 = [400, 280] \text{ Wm}^{-2}\) (hence a solar constant of 340 \text{ Wm}^{-2}), atmospheric absorption to \(b = 0.2\), and planetary albedo to \([a_1, a_2] = [0.12, 0.48]\) (hence, a global albedo of 0.3), to yield global forcing \(\bar{q} = 170 \text{ Wm}^{-2}\). The latter is balanced by the global convective flux \(\bar{f}_c = 100 \text{ Wm}^{-2}\) and surface LW flux \(\bar{f}_{LW} = 70 \text{ Wm}^{-2}\) (17). Additionally, setting \(T = 20 ^\circ\text{C}\) [59], the model solutions (4) and (6) yield \(T_2 = 12 ^\circ\text{C}\) and \(T_3 = 8 ^\circ\text{C}\).

Together with the variations derived in Sections 3 and 4, I calculate the temperature and ice cover at the thresholds. Connecting them by smooth curves and incorporating more rapid cooling at the thresholds, I produce the modeled time series shown in Figure 3 (thin lines). I have varied global sensitivity (5) over its outer range, between 0.5 (thin solid line) and 1.5 (thin dashed line) \(^\circ\text{C} (\text{Wm}^{-2})^{-1}\) [9], so solution spans (shaded portions) underscore their considerable uncertainty, but the general features remain robust. As these time series represent the prognosed cooling driven by the uplift-induced ice cover, I offer the following interpretation of the Pliocene climate.

Since the uplift of the Tibetan and American West highlands is largely confined to the subpolar zone (that is, north of 30°N), the resulting albedo increase has a greater effect on differential than global forcing—by three-fold (10)—leading to stronger subpolar than global cooling (13). The cooling is further augmented for the Arctic Ocean due to blockage of the incoming ocean heat, resulting in Arctic cooling that could double the rate of global cooling (14). It is significant to note that this Arctic amplification is comparable to that caused by anthropogenic warming [60] even without ice-covered Arctic Ocean and the accompanying ice-albedo feedback.

When the Arctic temperature falls below the GMT of 5 \(^\circ\text{C}\) around 3.5 Ma (the threshold labeled GIS), the pre-existing ice sheet over East Greenland would expand to West Greenland, descend the slope to the sea level, and advance toward the GMT. The glacial advance, however, is cut short by the shoreline to calve into the Nordic Seas. The rapid ice expansion on the millennial timescale may account for the observed abrupt cooling, which would be accompanied by minor IRD events, whose source has been traced to Greenland [3].
The slow uplift and attendant ice-albedo feedback continues to cool the climate, and when the Arctic Ocean is cooled to the freezing point (the threshold labeled NHG), perennial ice would grow to cover the Arctic Ocean with only a small open area maintained by nonzero ocean heat entering through the Fram Strait, as is the current case. The ice-covered Arctic Ocean would expel the snowline to the surrounding lowlands to allow a foothold of positive ice mass balance when perturbed, which would then grow the polar cap toward the GMT. The difference from the GIS threshold is that without seeding by the mountain glaciers, the polar cap can be initiated only when the sea level has cooled to the freezing point (not just the GMT), which represents a more stringent requirement. This is the reason that there is no widespread glaciation over the circum-Arctic lands until NHG [61], which must thus be distinguished generically from the GIS threshold. Again, the rapidity of the ice expansion causes the more abrupt cooling around 2.7 Ma, and the polar cap expanding southward over the contiguous lands would calve into both the Arctic Ocean and Nordic Seas to dispense massive IRD, the defining marker of NHG [5]. Contrary to earlier analysis [62], our model suggests that the Arctic Ocean should be largely free of perennial ice until NHG, although its observational evidence remains debated [63,64].

As expected, smaller global sensitivity (thin solid lines) would reduce the global and subpolar cooling relative to Arctic cooling, so a more expansive ice cover is needed for attaining NHG, but even the low end of this sensitivity requires only 17% of the cold-box area to be uplifted. If we assume that the Tibetan and American West highlands comprise 30% of the subpolar land [19], which occupies 60% of the cold box, the uplifted area would be 18% of the cold-box area (solid square), which hence is sufficient to cause NHG and even more so for greater global sensitivity. In addition, as we noted earlier, the cooling would lower the snowline to expose lesser highlands to glaciation, including the West Greenland ice sheet formed during the first threshold, further augmenting the uplift effect. With the above, I reckon that the Neogene uplift and its boost of the planetary albedo are sufficient to cool the Pliocene to NHG.

As discussed in Section 4, the generation of a polar cap over contiguous northern lands during NHG would activate the ice-albedo feedback to implement precession forcing and amplify obliquity forcing. Together with the augmented ice volume response, the glacial cycles would exhibit a stronger $\delta^{18}O$ signal, as observed [65]. Supported by the physics delineated above, I reckon, therefore, that the observed amplification of glacial cycles is the outcome rather than the driver of NHG. Although the reverse causality has been suggested [56], no physical mechanism has been advanced. Nor can it be attributed to changing orbital parameters, which simply do not vary on such timescales.

With continuing cooling by the uplift of circum-Arctic lands [66] and lowering of the global convective flux due to drier air, a threshold (labeled MPT) would be crossed when the glacial state would last through low eccentricity to enable LIS that defines the ice age. Again, ice expansion occurs over short mass balance timescale, so the threshold may account for the more abrupt cooling around 1 Ma, which represents MPT from obliquity-dominated glacial cycles to 100 ky ice-age cycles paced by eccentricity [52]. Ref. [36] provides a detailed discussion of MPT, and Figure 3 offers a tangible constraint on its timing, namely, when the global temperature cools to below about 10 °C.

6. Discussion

Although climate models are widely used to simulate past climate, their limitations should not be overlooked. In particular, the ocean is teeming with eddies, whose effect is incorporated through sub-grid parameterization finely tuned to yield the current climate [67], which necessarily downgrades their prognostic utility for a vastly different paleoclimate. The sensitivity can be lessened if climate models have resolved eddies, which is computationally challenging and remains to be attempted in paleoclimate studies. Our box model, while extremely crude, has nonetheless accounted for eddies by a closure based on probability laws of the nonequilibrium thermodynamics, namely the MEP—a generalized second law—and its deduced differential temperature and meridional overturning circulation,
which involve no free parameters, are commensurate with those observed in support of the MEP. Without resolving eddies to entail MEP, climate models may not be used to arbitrate our deductions, which can only be tested against observation. Given the crudeness of our box model and the uncertainty in observations, this comparison is limited to broad qualitative features, as depicted in Figure 3. To the degree that the modelled time series resemble the observed ones, including three abrupt cooling episodes, it is suggested that the model has captured the governing physics of the Pliocene cooling.

A critical parameter of our model is the glacial marking temperature (GMT) when net accumulation above the snowline is depleted by ablation down the slope. I have provided a derivation of this temperature in Appendix B, which turns out to be primarily an intrinsic property of the ice sheet, and the deduced value of ~5 °C is consistent with the present observation of the Greenland ice sheet [43] (their Figure 8). A fixed GMT sets the vertical drop (about 830 m) of glacial line from snowline, so the two would move in tandem by a changing climate, which moreover is through temperature, not moisture transport. The short-circuiting of the latter is because only the one crossing the snowline of freezing point is of relevance, which is thus insulated from the SST.

7. Conclusions

Given the many questions raised about previous proposals, the tectonic origin of the Pliocene cooling remains unresolved. Taking note of the Neogene uplift of the Tibetan highland, I posit that its icing above the snowline, hence augmented albedo, may initiate the Pliocene cooling. To test this hypothesis and more generally to explain the time evolution of the Pliocene climate, I formulate a minimal warm/cold/Arctic box model that is thermodynamically closed to prognose the Pliocene SST (same as the summer SAT on account of the well-validated convective bound). After the initial trigger, the ice-albedo feedback would continue the cooling, and three thermal thresholds are discerned: (1) when the Arctic SST falls below the GMT, the ice sheet over the East Greenland would descend to the sea level, accompanied by episodic IRD events; (2) when the Arctic SST cools further to the freezing point, the ice cap would form and expand on circum-Arctic lowlands, as manifested in massive IRD deposition that defines NHG; (3) when global convective flux is sufficiently lowered by the global cooling, the glacial state may last through low eccentricity to cause transition from obliquity- to eccentricity-dominated glacial cycles.

Given the short timescale governing the ice mass balance, the above thermal thresholds should appear as more rapid cooling punctuating the gradual Pliocene cooling, as indeed seen in the observed subpolar SAT. Fixing timing of the modelled thresholds to the observed ones around 3.5, 2.7, and 1 Ma, I produce a temporal evolution of the Pliocene climate as well as its driving albedo change. Since the latter can be accommodated by the observed uplift and growing ice cover, the model provides a causal explanation of the Pliocene cooling via the interplay of uplift-induced albedo, temperature, and glacial regimes. This supports our hypothesis that the Neogene uplift of the Tibetan highland is the tectonic origin of the Pliocene cooling and NHG.

**Funding:** This research received no external funding.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** No new data were created or analyzed in this study. Data sharing is not applicable to this article.

**Acknowledgments:** I thank anonymous reviewers for comments that have improved the paper.

**Conflicts of Interest:** The author declares no conflicts of interest.
Appendix A. Symbols

- $a$: Ice albedo ($= 0.9$)
- $a_i$: Planetary albedo of box-i
- $A_b$: Ablation
- $A_c$: Accumulation
- $b$: Atmospheric absorption ($= 0.2$)
- $B$: Lateral Bowen ratio ($= 4$)
- $F_a$: Atmospheric energy transport
- $F_l$: Latent heat transport
- $h$: Ice height
- $h_c$: Glacial line height
- $h_s$: Snowline height
- $\Delta h$: Height of snowline above the glacial line
- $i$: Ice cover
- $l$: Meridional span of plateau ($= 10^3$ km)
- $L_a$: Latent heat of sublimation ($= 2.84 \times 10^6$ J Kg$^{-1}$)
- $m$: Melt rate ($= \lambda T$)
- $p$: Annual precipitation rate
- $q_i^r$: Incoming SW flux of box $i$
- $q_i^s$: Absorbed SW flux of box $i$
- $\bar{q}$: Global-mean absorbed SW flux
- $\bar{q}_c$: Convective flux
- $\bar{q}_r$: Net radiative flux through tropopause
- $\bar{q}_{OLR}$: Outgoing LW radiation
- $q_{SW}^i$: Incoming SW flux
- $\Delta r$: Excess reflectance of ice over land
- $s$: Slope ($= 10^{-2}$)
- $\bar{\sigma}$: Global sensitivity ($= 1$ C(Wm$^{-2}$)$^{-1}$)
- $T_i$: SST (same as summer SAT) of box $i$
- $T_{a,i}$: SAT of box $i$
- $y$: Latitudinal distance $= \sin($latitude$)$
- $\alpha$: Air-sea exchange coefficient ($= 10$ Wm$^{-2}$C$^{-1}$ [29])
- $\gamma$: Lapse rate ($= 6$ C km$^{-1}$)
- $\lambda$: Melt rate per degree temperature increase ($= 0.8$ my$^{-1}$C$^{-1}$)
- $\mu$: Moisture content parameter ($= 2 \times 10^{-3}$ m$^3$ W$^{-1}$y$^{-1}$)
- $\rho_w$: Water density ($= 10^3$ Kg m$^{-3}$)
- $\sigma_a$: Atmospheric entropy production
- $\sigma_o$: Oceanic entropy production

Appendix B. Glacial Marking Temperature (GMT)

I provide here a derivation of the marking temperature of the glacial line, a key parameter of the model. As sketched in Figure A1 (see for example [43], their Figure 2a, July panel), I consider the north–south section of an idealized plateau when the snowline (thick solid line) is lower than the plateau height, so the moisture transport crossing the snowline would all manifest in the accumulation (no need to distinguish snowline and equilibrium line for our purpose). This moisture transport is seen later to be dominated by the summer season, so moisture and energy transports considered below pertain to averages of the summer-half of the year (referred to as ‘summer’ for short), which moreover are assumed to be in balance—neglecting seasonal storage of air and ice on account of their small thermal inertia. Assuming north–south symmetry of the ice sheet over the plateau, only half of the accumulation would descend the southern slope until it is depleted by ablation, which defines the glacial line (thick dashed line). The aim is to determine its temperature, referred to as the GMT.
I first link the moisture transport to the energy transport, which is assumed to be dominated by eddies as katabatic-type overturning circulation is appreciable only in winter and confined to the slope [68] and since quasi-geostrophic eddies do not transport potential energy [37] (their Section 13.3.3.2), the energy transport consists mainly of sensible- and latent-heat, and with their ratio denoted by $B$ (referred to as ‘lateral Bowen ratio’), the moisture transport (same as the accumulation $A_c$) can be linked to the latent-heat ($F_L$) hence the atmospheric energy ($F_a$) transports via

$$A_c = \frac{\rho w L_s}{\rho w L_s (1 + B)} F_a \equiv \mu F_a$$  \hspace{1cm} (A1)

where

$$\mu \equiv \left[ \rho w L_s (1 + B) \right]^{-1}$$  \hspace{1cm} (A2)

is a ‘moisture-content’ parameter with $\rho_w$ being the water density and $L_s$, the latent heat of sublimation. Since the vertical vapor profile tracks the saturated one, [69] has derived an analytical expression of the lateral Bowen ratio as a function only of the surface temperature $T$ and its saturation vapor pressure $e_s$ (his Equation (23)). Applying the Clausius-Clapeyron (C-C) relation approximated by [70] (their Equation (25))

$$e_s = 6.1 e^{0.074T} \text{mb},$$  \hspace{1cm} (A3)

we derive

$$B = (3.86 + 0.014 T) e^{-0.074 T}$$  \hspace{1cm} (A4)

Over the snowline defined by the freezing point, (A4) yields $B = 3.86$, which agrees with its observed value [69] (his Figure 4), yielding $\mu = 2.3 \times 10^{-3}$ m$^3$W$^{-1}$y$^{-1}$.

I next link the atmospheric energy transport to the radiative forcing, which states

$$F_a = l q_r,$$  \hspace{1cm} (A5)
where \( l \) is the meridional span of the plateau and \( q_r \) is the radiative cooling through the tropopause, given by

\[
q_r = q_{\text{OLR}} - q_{\text{SW}}(1 - a),
\]

(A6)

where \( q_{\text{OLR}} \) is the outgoing LW radiation, \( q_{\text{SW}} \) the incoming SW flux, and \( a \) the ice albedo. Combining (A1) and (A3), the annual (hence the 1/2 factor below) precipitation rate over the plateau is

\[
p = A_c/(2l) \quad (A7)
\]

\[
= \mu q_r/2. \quad (A8)
\]

For a cursory check of the Greenland ice sheet, which rises to 2 km above the snowline situated at 1 km, the mean plateau temperature is \(-6^\circ\text{C}\), so \( q_{\text{OLR}} = 224 \text{ Wm}^{-2} \) [71] (his Equation (23) and Figure 2). Setting \( q_{\text{SW}} = 400 \text{ Wm}^{-2} \) and \( a = 0.9 \), we have \( q_r = 184 \text{ Wm}^{-2} \), hence, the annual precipitation rate is \( p = 0.21 \text{ m/y} \), which is commensurate with the observed one [72]. Given the simple expression (A8), this agreement underscores the strong constraint placed on the accumulation by energy balance and C–C relation—irrespective of the vagaries of the atmospheric motion that carries moisture.

To assess the neglected winter precipitation, I set the winter temperature to be 20 \( ^\circ\text{C} \) cooler than summer [43], so the mean plateau temperature is \(-26^\circ\text{C} \) to yield \( q_{\text{OLR}} = 180 \text{ Wm}^{-2} \) hence a radiative cooling of \( q_r = 180 \text{ Wm}^{-2} \) neglecting the SW flux in winter. It is noted that this radiative cooling is comparable to that of the summer, in contrast to that over the ocean when the poleward energy transport is typically smaller in summer [37] (their Figure 13.11). This is because the ice albedo strongly curtails the SW heating in summer to boost the radiative cooling hence the poleward energy transport. With the snowline now at \(-20^\circ\text{C}\), (A.4) implies \( B = 15.7 \) hence \( \mu = 0.66 \times 10^{-3} \text{ m}^3\text{W}^{-1}\text{y}^{-1} \), which yields annual precipitation rate of \( p = 0.06 \text{ m/y} \). It is seen that although summer and winter energy transports are comparable, the drier air in winter renders moisture transport hence precipitation that is only 1/3 of the summer, justifying its neglect in GMT.

To derive the ablation \( A_b \), I use the melt rate \( m \) based on the positive-degree day method [73]

\[
m = \lambda T, \quad (A9)
\]

where \( \lambda = 0.8 \text{ my}^{-1}\text{C}^{-1} \) [36], and assume a constant slope \( s = 10^{-2} \) to derive

\[
A_b = \int_{x_s}^{x} m \, dx
= -\frac{\lambda}{s} \int_{h_s}^{h_s-\Delta h} T \, dh
= -\frac{\lambda \gamma}{s} \int_{h_s}^{h_s-\Delta h} (h_s - h) \, dh
= \frac{\lambda \gamma}{2s} \Delta h^2
= \frac{\lambda \gamma}{2s} \frac{T_s^2}{\gamma},
\]

(A10)

where \( x \) is the southward distance, subscripted \( s \) and \( g \) for snowline and glacial line, respectively; \( h_s \), the snowline height; \( \Delta h \), its elevation above the glacial line; and \( T_s = \gamma \Delta h \) is the GMT. Since only half the accumulation descends the southern slope, the mass balance states

\[
A_c/2 = A_b, \quad (A11)
\]

Substituting from (A1), (A5), (A10) and applying the foregoing parameter values (see also Appendix A) yields the GMT

\[
T_g = \left( \frac{ls \lambda q_r \mu}{\lambda} \right)^{1/2} \approx 5.3^\circ\text{C}. \quad (A12)
\]
This deduced GMT is consistent with the observed Greenland ice sheet [43] (their Figure 8). Although the parameters of (A12) are uncertain, the sensitivity is dampened by the $\frac{1}{2}$ power dependence. More significantly, it underscores the primary control of the GMT by ice-sheet properties, so changing climate would displace snow- and glacial lines in tandem, which, however, is through temperature, not moisture transport. The short-circuiting of the moisture transport occurs because only the one crossing the snowline (defined by the freezing point) causes the accumulation, which is thus insulated from the SST.

References


18. Kerrick, D.M.; Caldeira, K. Was the Himalayan orogen a climatically significant coupled source and sink for atmospheric CO$_2$ during the Cenozoic? *Earth Planet Sci. Lett.* 1999, 173, 195–203. [CrossRef]


21. Milankovitch, M. *Canon of Insolation and the Ice-Age Problem*; Royal Serbian Academy Publication: Belgrade, Serbia, 1941; Volume 132.


56. Knies, J.; Matthiessen, J.; Vogt, C.; Stein, R. Evidence of ‘Mid-Pliocene (~3 Ma) global warmth’ in the eastern Arctic Ocean and implications for the Svalbard/Barents Sea ice sheet during the late Pliocene and early Pleistocene (~3–1.7 Ma). *Boreas* 2002, 31, 82–93. [CrossRef]


