***Nature Communications* *Review***

**Integrating Pliocene Geological Archives and Climate Models: Progress, Challenges and Future Direction**

**Alan M. Haywood1 and Harry J. Dowsett2**

1School of Earth and Environment, University of Leeds, Woodhouse Lane, Leeds, West Yorkshire, LS2 9JT, UK: earamh@leeds.ac.uk

2Eastern Geology and Paleoclimate Science Center, US Geological Survey, 12201 Sunrise Valley Drive, Reston, Virginia, 20192, USA: hdowsett@usgs.gov

**Summary paragraph** (100) words**:**

The Pliocene is recognised as a priority for environmental reconstruction and climate modelling. Pliocene warm intervals provide documentation and better understanding of a warmer-than-present world, and the predictive ability of climate models. The relationship between environmental reconstruction and modelling is central to the Pliocene’s success in palaeoclimatology. Geologists and climate modellers are highlighting critical features of Pliocene climate/environments such as polar amplification, ice sheet and sea level change, and altered patterns of surface temperature and precipitation. The synergy between both communities is identifying challenges and methodological improvements to meet those challenges in order to enhance our understanding of global change.

**1. Introduction** (797 words, 1 figure)

The latest Intergovernmental Panel on Climate Change Assessment Report (AR5: IPCC, 2013) states that “warming of the climate system is unequivocal, and since the 1950s, many of the observed changes are unprecedented over decades to millennia. The atmosphere and oceans have warmed, the amounts of snow and ice have diminished, sea level has risen, and the concentrations of greenhouse gases have increased”. Records of atmospheric carbon dioxide (CO2) stored within bubbles of air trapped within ice cores reveal that current concentrations of atmospheric CO2, as well as other greenhouses gasses such as Methane (CH4) and Nitrogen Dioxide (N2O), are unprecedented over the last 800,000 years of Earth history (IPCC, 2013).

Earth’s climate and environmental evolution is preserved within an exciting and diverse array of geological archives (Vaughan, 2007). These archives indicate that climate can respond rapidly to different forcings (such as CO2 concentration). Earth’s geological history provides the scientific community with natural laboratories in which to understand Earth system processes that operate over timescales longer than any observational or instrumental record, and are fundamental to our understanding and ability to accurately predict future climate and environmental change (Valdes, 2011).

The Pliocene epoch (5.33 to 2.58 Ma) has become a highly valued geological target in which to explore climate processes in a warmer world. During warm intervals of the Pliocene, atmospheric CO2 concentration is estimated to have ranged between 350 to 450 ppmv (e.g. Raymo et al., 1996; Kürschner et al., 1996; Seki et al., 2010; Pagani et al., 2010; Bartoli et al., 2011; Badger et al., 2013). This is in contrast to the known pre-industrial concentration of 280 ppmv and the seasonally adjusted 2013 annual concentration of 396.52 ppmv (http://scrippsco2.ucsd.edu/data/atmospheric\_co2.html). With respect to the pre-industrial era, Pliocene surface temperatures over land and oceans were elevated (Dowsett et al., 2012; Salzmann et al., 2013), and climate model estimates indicate that the global annual mean surface temperature was 2.70 to 4.05°C higher (Haywood et al., 2013a). The hydrological cycle was enhanced (Haywood et al., 2013a), ice sheets were smaller (Cook et al., 2014; Dolan et al., 2011), sea-level was higher (Dowsett and Cronin, 1990; Miller et al., 2012), forest cover was expanded (Salzmann et al., 2008) and arid deserts were contracted (Salzmann et al., 2008). Many parallels can easily be drawn between warm intervals of the Pliocene and observed trends in modern climate, as well as in future climate predictions by the similarity of Pliocene continental configuration, mountain elevation and ocean bathymetry to today (Dowsett et al., 2010). Finally, many Pliocene terrestrial and marine flora and fauna are extant, increasing the validity of approaches that attempt to reconstruct Pliocene surface conditions using modern observed relationships between climate and biogeography (Dowsett et al., 1996; Thompson and Fleming, 1996).

Scientific endeavors focused on the entire Pliocene with new high-resolution records documenting temperature variability is appearing in the literature on a regular basis (e.g. Wara et al., 2005; Lawrence et al., 2010; Brigham-Grette et al., 2013). Maximum data coverage, however, has been achieved through a concentrated effort in environmental reconstruction for the mid-Pliocene Warm Period (mPWP: 3.3 to 3.0 Ma; Dowsett et al., 1999; Dowsett, 2007a) (**Figure 1**). The mPWP is the last interval of generally sustained warmth prior to the intensification of Northern Hemisphere glaciation approximately 2.75 Ma (Dowsett et al., 2010). Over the last 25 years, scientists have systematically reconstructed ocean temperatures, vegetation cover and ice extent producing digital data sets specifically designed for integration with climate models (Dowsett et al., 2010), thus enabling models to be used to explore the nature of Pliocene climates on a global scale (Chandler et al., 1994). The Pliocene, and specifically the mPWP, represents one of the best examples currently available of successful synergy between the geological and climate modelling communities. The benefits derived from this synergy are clearly manifested in documenting, but also in understanding, Pliocene climate and environmental change. Over the last 25 years, the relationship between models and data has evolved and has become increasingly intertwined and complex. The increasingly sophisticated nature of geological reconstruction and climate modelling techniques have driven a continuous process of innovation in combined data and modelling approaches to understanding the mPWP and what it may be able to tell us about future global change (Chandler et al., 1994; Sloan et al., 1996; Haywood et al., 2000a/b; Haywood et al., 2013a). This review documents the evolution of geological and modelling efforts to understand the mPWP, and the way in which the data and modelling communities are able to inform each other on important issues associated with climate variability, processes driving local, regional and global climate change, and uncertainties in palaeoenvironmental reconstruction and climate modelling. It also examines exciting recent developments in Pliocene science that are likely to lead to step changes in our knowledge in the future.

**2. Brief historical perspective of Pliocene palaeoenvironmental reconstruction** (1766 words, 1 figure)

The palaeoclimate of the Pliocene has been documented in the literature since Lyell (1845) compared the resemblance of fossil shell assemblages in Virginia, USA, to those found in Suffolk, UK. He commented on the similarity between the faunas occurring on opposite sides of the Atlantic while noting their latitudinal differences. This was, in effect, an early observation of the palaeo-position and impact of the Pliocene Gulf Stream and North Atlantic Drift.

Zubakov and Borzenkova (1988) pioneered the earliest efforts towards the construction of Pliocene palaeoclimate conditions regionally based upon the palaeontology of more than 20 continental sections and available marine core sequences. From these data they reconstructed 30 Miocene and Pliocene ‘superclimathems’ which they defined as cycles of 100,000 to 300,000 years with amplitudes of at least 4 to 5°C. They were the first to propose that the climate of their defined Pliocene Optimum (4.3-3.3 Ma) could be considered a past analogue of the mid-21st century when atmospheric CO2 concentrations would reach double their pre-industrial values.

With the then generally accepted knowledge that the Pliocene was a time of warm, equitable climate, and growing concern over potential impacts of future global warming as a backdrop, the need for a more precise and less anecdotal assessment of Pliocene climatic conditions became clear. In 1988 the U.S. Geological Survey endeavored to reconstruct the Pliocene palaeoenvironment of the planet via Pliocene Research, Interpretation and Synoptic Mapping (PRISM). Borrowing heavily from lessons learned from earlier efforts to reconstruct the surface of the Ice-Age Earth (CLIMAP, 1976), PRISM and its international collaborators systematically developed a large-scale data collection project that has grown in size and scope over the past 25 years (**BOX 1**). PRISM is, until now, the only *synoptic* reconstruction of the Pliocene. Data are produced from a global distribution of localities, but work is concentrated on a focused stratigraphic interval. Considerable work by others has gone into developing long time series that intersect the mPWP, without which our ability to understand the dynamic development and evolution of Pliocene palaeoclimate would be constrained (see also next section). These approaches (time slice and time series) are not mutually exclusive. Rather, each benefits from the perspective only available through the other.

Disproportionate amounts of marine versus terrestrial data are available for Pliocene palaeoenvironmental reconstruction. Work in the terrestrial realm is dependent upon the heterogeneous distribution of localities where palaeoclimate signal carriers are preserved. Those outcrops and cores with suitable Pliocene chronology, or that contain continuous sequences, are few in comparison to the quantity of cores containing Pliocene marine sediments retrieved by the International Ocean Discovery Program (IODP) and its predecessors. Terrestrial records from near shore marine cores allow for high-resolution marine-continental correlation of Pliocene palaeoclimate records (e.g. DeMenocal, 2004). In addition, Pliocene material retrieved from remote locations, like Lake El´gygytgyn in Siberia, via the International Continental Drilling Program (ICDP) is proving invaluable to our understanding of Arctic continental climate variability (Brigham-Grette et al., 2013).

The tools used by the Pliocene palaeoclimate community have changed over time, and researchers continue to use the most sophisticated techniques available for reconstruction (**Figure 2**). Proxies have been developed to estimate different aspects of the environment, but surface ocean and land temperatures are by far the primary data generated. These proxy techniques include quantitative analysis of faunal and floral assemblages, stable isotopic composition of carbonates and biomarkers. All have strengths and limitations, but each plays an important role in our conceptual understanding of the Pliocene Earth.

Stable isotopes of oxygen have been used to estimate palaeotemperatures, and indirectly ice volume, since the pioneering work of Emiliani (1955). Over time they have become the multi-tool of palaeoceanographic inquest. Shackleton et al. (1995) generated the first long benthic δ18O time series spanning the Pliocene, placing limits on potential changes in the cryosphere, the stability of Antarctic ice, and sea level. The Central American Seaway (CAS) may have played a significant role in the initiation of Northern Hemisphere glaciation, biotic evolution, and surface hydrography of the North Atlantic Ocean.

The timing of fossil land mammal migration between North and South America, and concomitant evolutionary divergence of marine faunas on opposite sides of the Isthmus of Panama, suggest that tectonic, oceanographic, biotic and climatic systems were intricately coupled during the Pliocene (Coates et al., 1992). Another seminal work based upon δ18O time series led to the now generally accepted date for this closing of the Central American Seaway (CAS) and initiation of faunal migrations (Haug and Tiedemann, 1998; Haug et al., 2001). These high-resolution records document increase in Caribbean salinity as a result of the shoaling seaway (**Figure 2**).

Oxygen isotopes have played an even larger role in Pliocene research; the development of an isotopic composite standard reference section (LR04; Lisiecki and Raymo, 2005) is possibly the most unifying development to date in pre-Pleistocene palaeoceanography. With the LR04 composite, researchers are able to correlate remote sequences to a standardized section tied to orbital chronology. This provides the community a high-resolution stratigraphic framework within which the temporal and spatial aspects of Pliocene climate change can be analysed.

Primary temperature proxies commonly utilized for Pliocene reconstruction include: Mg/Ca palaeothermometry, the alkenone unsaturation temperature proxy (), the TetraEther index (TEX86H) and quantitative analysis of faunal and floral assemblages. These palaeotemperature proxies measure different aspects of temperature by sampling the marine environment at various spatial and temporal resolutions. Estimates are complicated by effects unique to different signal carriers and methods. For example, the Mg/Ca method depends upon assumptions about the composition of water at the time calcite was precipitated and can be complicated by post-depositional processes like dissolution. Faunal assemblage techniques require the assumption of stationarity of ecological preferences and have upper and lower limits based upon calibration to present day conditions. , based upon extraction of organic molecules (ketones synthesized by haptophyte algae) from sediment, is not affected by seawater chemistry but has an upper limit of ~28°C. TEX86H (based on the relative distribution of membrane lipids produced by Crenarchaeota), like , is not directly affected by seawater composition, has an upper limit close to 38°C, but can sometimes record subsurface conditions and be complicated by terrestrial input (Lea et al., 1999; Barker and Elderfield, 2002; Rosenthal et al., 2000; Imbrie and Kipp, 1971; Dowsett and Poore, 1990a/b; Sachs et al., 2013; Volkman et al., 1980; Conte et al., 1994; Prahl et al., 1988; Muller et al., 1998). The use of multiple proxies is of great benefit to gaining more robust and detailed estimates of the palaeoenvironment as long as the relative limitations of various techniques are recognised.

The concept of permanent El Niño–like conditions and thermal stability of the western equatorial Pacific (WEP) warm pool through a tropical thermostat mechanism has been a driver of Pliocene palaeoceanographic research for many years and is a good illustration of the application of multiple proxies. Wara et al. (2005) developed Mg/Ca-based sea surface temperature (SST) records from calcareous tests of surface dwelling planktonic foraminifers. These time series span the Pliocene to Recent at Sites 847 in the eastern equatorial Pacific (EEP) and 806 in the WEP (**Figure 2**). The records show a small surface temperature gradient across the equatorial Pacific, throughout the Pliocene, much like a modern El Niño event with warmer than average SST in the EEP. This led to the hypothesis of *Permanent El Niño* during the Pliocene, with concomitant warming in the EEP suggesting weak zonal atmospheric circulation.

The Site 806 Mg/Ca record suggests stable temperatures while climate model simulations with increased atmospheric CO2 call for 2° to 3°C warming in the WEP (Haywood et al., 2013a). Dowsett (2007b) used quantitative faunal assemblage techniques to estimate SST and found no compelling evidence that the WEP was different than present day. Since faunal techniques (calibrated to present day) have a maximum limit of about 30°C, and index becomes fully saturated at 28°C, neither can document warming above modern warm pool conditions. However, some Late Pliocene assemblages at Site 806 do show small non-analogue increases in thermophillic taxa that could be explained by brief periods with SST in excess of 30°C.

More recently O’Brien et al. (2014) and Zhang et al. (2014) addressed the issue of stability of the tropical warm pools during the Pliocene by applying Mg/Ca, and TEX86H techniques. They found close agreement between and TEX86H where they could both be applied, giving confidence to the application of TEX86H to the warm pool regions. The TEX86H -based estimates of WEP SST were higher than previous studies based upon Mg/Ca or faunal assemblages. They concluded that previous WEP Mg/Ca temperatures were underestimated due to changes in the Mg/Ca of Pliocene seawater (see section 3 for more discussion of the Pliocene tropical Pacific and ENSO).

Arguably the longest and most complete record of Pliocene high latitude continental climate comes from Lake El’gygytgyn (Brigham-Grette, 2013), located in Northeast Arctic Russia in a basin formed ~3.6 Ma by a meteorite impact. The many environmental proxies from the Lake El’ gygytgyn core can be correlated to the LR04 stack and thus placed within the same chronostratigaraphic framework as marine core sequences (**Figure 2**). These records document polar amplification such as is seen in marine records (Dowsett et al., 2013) but that appears to be under-estimated in model simulations (Haywood et al., 2013a). This sequence documents mid-Pliocene summer temperatures 8°C warmer than present day and warmer than present day summers until ~2.2 Ma. This supports other estimates of strong Arctic warming during parts of the Pliocene (Ballantyne et al., 2010).

Atmospheric concentration of CO2 during the Pliocene is only partially constrained. A number of techniques exist to estimate CO2 (alkenones, B/Ca, δ11B, δ13C, leaf stomatal density), and the variability of the estimates is high (**Figure 2**). None of the present estimates suggest atmospheric concentrations much higher than present day (Raymo et al., 1996; Tripati et al., 2009; Pagani et al., 2010; Seki et al., 2010; Bartoli et al., 2011), making the cause of Pliocene warmth challenging to identify (Fedorov et al., 2013). It is increasingly recognised, however, that CO2 is just one agent of radiative forcing and that other greenhouses gases such as Methane (CH4) are also important but cannot be reconstructed at the current time. Furthermore, Unger and Yue (2014) have demonstrated the potential importance of atmospheric chemistry-climate feedbacks, as well as aerosols, in augmenting surface temperature warming derived from a given increase in atmospheric CO2. This represents a huge challenge to the scientific community to assess the efficacy of the increasingly complex simulations of additional Earth System components in climate models in the absence of any direct geological records (see section 3 for further discussion).

1. **The emergence of palaeoclimate model simulations for the Pliocene** (1484 words, 2 figures)

Even before the publication of global geological data sets capable of underpinning climate model simulations for the Pliocene, climate modellers were using emerging regional patterns of surface temperature change to construct conceptual models and proposing mechanisms and drivers of global and regional change. The modest CO2 increase reconstructed for the mPWP was calculated to be capable of producing a radiative forcing of approximately 2 Wm-2. This may have been sufficient to explain the warmth of the Pliocene globally, depending on the chosen climate sensitivity value, it could not explain, however, the regional changes in reconstructed SST (based on planktonic foraminifera) that initially displayed little or no warming relative to present day in the tropics (Dowsett et al., 1996; Dowsett et al., 1999). Therefore, an additional mechanism(s) working independently or combined with variations in CO2 concentrations in the atmosphere was thought to be required to drive and maintain the warmer conditions reconstructed for the mid-Pliocene Warm Period (Crowley, 1996).

A likely additional driver of Pliocene warmth receiving attention in the early part of the 1990’s, and still discussed today, was a change in the ocean heat transport. This initially geologically driven hypothesis (Dowsett et al., 1992) received further impetus from a conceptual modelling study published by Rind and Chandler (1991). However, Crowley (1991; 1996) pointed out a paradox in that a reduced latitudinal SST gradient implies potentially weaker atmospheric forcing of oceanic circulation, and hence weaker oceanic heat transport. Thus, the total benefit in terms of the 'global net' heat transport from enhanced ocean heat transport was lower than would otherwise be expected (Crowley, 1996). This early finding based upon a conceptual framework has recently been supported by the investigation of Pliocene ocean circulation and heat transport behavior in climate models (Zhang et al., 2013).

Early fully numerical simulations of the mPWP climate were generated using atmosphere-only General Circulation Models (AGCMs). These models required monthly SSTs and sea-ice, as well as vegetation cover, to be prescribed from the PRISM data set, and these were held constant for the entire length of the model simulation. Whilst this negated the possibility of examining atmosphere-ocean-vegetation feedbacks, it did provide a way to explore the response of the atmosphere, in isolation, to a given set of geological boundary conditions. The first of these simulations was conducted by Chandler et al. (1994) using the PRISM0 reconstruction and focused on responses in the Northern Hemisphere. The second study by Sloan et al. (1996) using the PRISM1 data set was the first simulation to be performed for the entire globe. Haywood et al. (2000) performed the second global simulation, the first to use the updated PRISM2 boundary conditions that incorporated a greater SST data density as well as revised reconstruction of the Antarctic Ice Sheet. The current PRISM data sets (PRISM3D) have been incorporated into a number of AGCMs (e.g. Bragg et al., 2012) as part of the Pliocene Model Intercomparison Project that will be discussed later (**Figures 3 and 4**).

AGCM simulations resulted in a global annual mean surface air temperature warming of 1.4 to 3.6°C. Warming was greatest at high latitudes; consequently, the equator to pole temperature gradient decreased. Surface air temperature increases were greatest in winter, as decreased snow and sea ice triggered a positive albedo temperature feedback. At low latitudes, temperatures were mostly unchanged except for cooling over Northern Africa (Chandler et al., 1994; Sloan et al., 1996; Haywood et al., 2000a/b). This model-predicted cooling was supported by palaeobotanical data and, in the simulations, was a response to a weakening and/or broadening of the Hadley Cell, which caused local subtropical clouds and evapotranspiration rates to increase. The hydrological cycle intensified, where regionally annual evaporation, rainfall, and soil moisture all increased (Chandler et al., 1994; Sloan et al., 1996; Haywood et al., 2000a/b).

AGCM simulations quickly evolved to include a simplified slab ocean capable of predicting SST and sea-ice change. These models were employed to explore the sensitivity of Pliocene climate to atmospheric CO2 concentration (Chandler unpublished data), as well as the sensitivity of climate predictions to orbital forcing (Haywood et al. 2002).

A major breakthrough was achieved through the use of fully coupled atmosphere-ocean GCMs (AOGCMs), where the response of the oceans was, for the first time, predicted. This milestone changed the relationship between models and marine data, and moved the use of marine data away from providing model boundary conditions and towards providing verification data for model evaluation. Haywood and Valdes (2004) explored the relative role of the atmosphere, oceans and cryosphere in contributing towards Pliocene warmth. In contrast to AGCM experiments, and available geological archives, surface temperatures warmed across the globe, including in the tropics. Analysis of the predicted ocean circulation indicated reduced outflow of Antarctic bottom water, a shallower depth for North Atlantic deep water formation and slightly weaker thermohaline circulation. Neither the oceans nor the atmosphere transported significantly more heat in the Pliocene scenario, and a significant driver of warmth was the reduced extent of high-latitude terrestrial ice sheets and sea ice cover resulting in a strong ice-albedo temperature feedback (Haywood and Valdes, 2004).

The progressive adoption of AOGCMs unquestionably represented a paradigm shift in Pliocene climate modeling and underpinned a host of further applications exploring, for example, the sensitivity of climate to greenhouse gas concentrations (Haywood et al., 2011a), modes of climate variability, such as the North Atlantic Oscillation (e.g. Hill et al., 2011) and ENSO (Bonham et al., 2009), the importance of ocean gateways and ocean bathymetry (Lunt et al., 2008), climate and longer term sensitivity (Lunt et al., 2010), the importance of uncertainties in model parameterization (Pope et al., 2011; Howell et al., 2014), as well as detailed studies exploring the importance of individual boundary conditions in driving Pliocene warmth and polar amplification (Lunt et al., 2012).

Since 2008 multiple AGCM and AOGCMs have been applied to the mPWP in a fully coordinated way through the Pliocene Model Intercomparison Project (PlioMIP). This, as much as the first use of AOGCMs, has served to reshape our thinking and understanding of Pliocene climate, and has rapidly driven forward the frontier of Pliocene climate modelling. For the first time, PlioMIP facilitated an exploration of robust model predictions - predictions that are reproduced by a number of models rather than a product of a single model (Haywood et al., 2013a). PlioMIP has applied new methodological rigor towards the understanding of large-scale features of climate and climate sensitivity (Haywood et al. 2013a), patterns of ocean circulation (Zhang et al., 2013), monsoon behavior (Zhang et al. 2013), and has assessed the ability of models to reproduce patterns of regional temperature change recorded by geological archives (Dowsett et al., 2012; Haywood et al., 2013a/b; Dowsett et al., 2013; Salzmann et al., 2013). As well as identifying potential weaknesses in model predictions, PlioMIP has served as an effective means to identify potential deficiencies in interpretation of geological data and inconsistencies in the methodology for data/model comparison, which will be discussed further in sections 5 and 6.

Increasingly, Earth System Models (models capable of including a variety of additional Earth system processes) are allowing entirely new components of the Pliocene climate system to be explored. However, they present their own unique challenges in terms of providing adequate geological constraints for the models and in evaluating the outputs of the models themselves. Earth System components have included interactive vegetation allowing vegetation/climate feedbacks to be incorporated within simulations (e.g. Haywood and Valdes, 2006). Ice sheet models are now being used to better understand the validity of ice sheet reconstructions used in climate models (e.g. Hill et al., 2007; Dolan et al. 2011; Dolan et al. 2014) and in addressing different estimates of Pliocene sea-level (Dowsett and Cronin, 1990; Wardlaw and Quinn, 1991; Brigham Grette and Carte, 1992; Kennett and Hodell, 1995; Miller et al., 2005, 2011; Sosdian and Rosenthal, 2009; Naish and Wilson, 2009; Dwyer and Chandler, 2009; Miller et al., 2012).

Finally, models of atmospheric chemistry are being used to simulate the effects of altered vegetation patterns on dust and aerosol emissions and the radiative forcing of such agents that have little or no expression within geological archives. For example, Unger and Yue (2014) calculated terrestrial ecosystem emissions and atmospheric chemical composition for the mPWP and the pre-industrial era. Tropospheric ozone and aerosol precursors from vegetation and wildfire were ~50% and ~100% higher in the Pliocene due to the spread of the tropical savanna and deciduous biomes. The chemistry-climate feedbacks contributed to a net global warming that is +30 to 250% of the carbon dioxide effect, and a net aerosol global cooling masked 15 to 100% of the carbon dioxide effect. Studies, such as Unger and Yue (2014), highlight the emergence of new unknowns in terms of atmospheric forcing in the Pliocene that are potentially more important than the currently known uncertainty in CO2 forcing, which from a modelling point of view can now be easily explored through CO2 sensitivity experiments.

1. **Data/Model Comparison** (1027 words, 1 figure)

The utilisation of AOGCMs in a Pliocene context rapidly transformed the role of geological archives of surface temperature from providing model boundary conditions to providing the means to independently evaluate model performance. This was an important step as it facilitated a new way to use the Pliocene to evaluate models that are used to predict future climate change (**Figure 5**). Geological archives of past temperature are normally produced as time series with the sampling density determining the time resolution of the record. As stated previously, there are now a number of long high-resolution temperature time series for the Pliocene that, theoretically, are ideal to evaluate model performance. In reality, however, the process of data/model comparison is complex, fraught with uncertainty and requires a full and open discussion of (a) the inherent strengths and weaknesses of climate models, (b) uncertainties in geological boundary conditions used within models to facilitate Pliocene simulations, and (c) uncertainties in the geological archives and their interpretation.

There is no *a priori* reason climate models should be expected to perform better in the Pliocene than in the modern. It is presently not possible to constrain the forcings required by models to simulate Pliocene climate to the same degree as for the modern. It also follows that estimated Pliocene surface temperatures will not attain the same accuracy as observations of modern temperature. Thus, it is logical to assume simulations of Pliocene climate may diverge from palaeoenvironmental *estimates* of surface temperature to a greater degree than simulations of modern climate diverge from *observations* of surface temperature. Given the inherent uncertainties already discussed in reconstructing and modelling the past, determining the reliability of the signal of Pliocene data/model discord becomes critical.

Due to limitations associated with computational efficiency, representing the coupled climate-ice sheet and carbon cycle in models, as well as inadequate scientific understanding of time varying boundary conditions and forcings throughout the Pliocene, all previous data/model comparisons have used data extracted from a time slab of the Pliocene (e.g. the PRISM time slab). Time slab data are then compared with equilibrium (snap-shot style) climate model simulations (e.g. Haywood and Valdes, 2004; Dowsett et al., 2011, 2012, 2013; Haywood et al., 2013a; Salzmann et al., 2013). Published data/model comparisons for the mPWP have:

* Identified many regions of agreement between data and model predictions, such as in the Southern Ocean (e.g. Haywood et al., 2013a; Fig. 5).
* Highlighted potential inconsistencies between data and models in the tropics, where models may predict surface temperatures that are too warm compared to data (Haywood et al., 2005; Dowsett and Robinson, 2009; Salzmann et al., 2008; Fig. 5).
* Suggested that model predicted SST change in the North Atlantic may be insufficient (Dowsett et al., 2012; Dowsett et al., 2013; Fig. 5).
* Indicated a potentially inadequate degree of polar amplification in models (Ballantyne et al., 2010; Salzmann et al., 2013; Dowsett et al., 2013; Fig. 5).

Furthermore, Fedorov et al. (2013) synthesized geochemical records of SST and suggested that compared to today, the early Pliocene climate had substantially lower meridional and zonal temperature gradients but similar maximum ocean temperatures. Using an Earth System Model, they attempted to demonstrate that none of the mechanisms currently proposed to explain Pliocene warmth can simultaneously reproduce the combined changes in meridional and zonal temperature gradients whilst at the same time maintaining similar to modern maximum ocean temperatures. The suggested cause was a combination of several dynamical feedbacks underestimated in the models at present, such as those related to ocean mixing and cloud albedo. Whilst one of the fundamental drivers of a reduced meridional temperature gradient was omitted in the study of Fedorov et al. (2013), specifically the specification of smaller ice sheets in the model simulations of the Pliocene, it is intriguing that the geological record is identifying possible structural failings in climate models.

At present a detached assessment of the suggested model shortcomings against uncertainties in the interpretation of surface temperature records from geological data can only conclude that the true implications of these patterns of data/model discord are uncertain. There are several reasons for this:

* PlioMIP has shown a substantial spread in model predictions. Therefore, previous studies that have analysed the performance of a single model and have suggested that Pliocene data/model comparison is able to demonstrate a structural weakness in climate models are too simplistic (Haywood et al., 2013a).
* Forcings and model experimental design are at best incomplete, or at worst incorrect. This covers an uncertainty in the agents of radiative forcing (including atmospheric greenhouse gas composition, chemistry/climate feedbacks: Unger and Yue 2014), ice sheet configuration (Dolan et al., 2014) and the likelihood that similarities drawn between modern and Pliocene topography, drainage and bathymetry are overstated.
* The interpretation of surface temperature patterns derived from geological archives are, in certain circumstances, still evolving (e.g. Zhang et al., 2014; O’Brien et al., 2014).

The current debate surrounding tropical Pacific Ocean temperatures is a key illustration of the last point. The foundation of the Fedorov et al. (2013) case is built upon the premise that maximum SSTs in the western Pacific warm pool during the Pliocene do not exceed modern temperatures. This is essential to maintain a lower tropical Pacific SST gradient resulting from elevated SSTs in the east and constant SSTs in the west. Specific proxies are limited to recording temperatures in the WEP warm pool that do not exceed the modern maximum temperature (see section 2). Mg/Ca-based temperature reconstructions that retain sensitivity at temperatures higher than modern, initially indicated no increase in tropical Pacific SSTs in the warm pool, supporting the Fedorov et al. (2013) argument. Recent revaluations of the Mg/Ca data to incorporate the effects of changing concentrations of Mg/Ca in the water column through time, however, as well as the application of a newer proxy (TEX86H), have indicated that surface ocean temperatures in the WEP warm pool may have been 1 to 2°C higher than modern. If true, this effectively resolves the mismatch between tropical Pacific Ocean temperature data and the PlioMIP ensemble (O’Brien et al., 2014) and critically removes the requirement on the models to predict the complete structural change in zonal and meridional SST gradients suggested by Fedorov et al. (2013).

1. **Outlook** (800 words, 1 figure)

The temporal resolution of the original PRISM marine reconstruction was limited by the available magnetobiochronology to a 300,000 year interval that could confidently be correlated on a global scale (Berggren et al., 1985; Dowsett and Poore, 1991). Within this time slab, warm phases of climate were averaged. This inadvertently created a situation where non-coeval conditions were smoothed and compared, and a super-equilibrium climate was created that may never have existed. The advent of a unifying chronology that is both applicable to select existing sites and the standard for new deep-sea sequences has had a major impact on the PRISM reconstruction and Pliocene paleoclimatology in general (Lisiecki and Raymo, 2005; Dowsett et al., 2013). Millennial-scale and even sub millennial-scale resolution now allows for development of short sequences that can be correlated to a single precession cycle (e.g. Herbert et al., 2015).

Enhanced constraints on age models and chronology have coincided with climate model-based studies that demonstrate that climate variability during the Pliocene has been incompletely sampled (Prescott et al., 2014). This realization opens a new chapter in the relationship between geological archives and climate models for the Pliocene; a relationship where climate modellers work in equal partnership with the geological community to 1) elucidate more relevant scientific questions, 2) identify new intervals in the Pliocene to study, and 3) develop new approaches to better identify and reduce uncertainties in reconstructing/modelling the Pliocene (Haywood et al., 2013b).

Understanding the importance of orbital forcing on Pliocene climate variability is vital to moving climate models and geological data forward. Whilst the benthic oxygen isotope record implies a reduced feedback between orbital forcing and ice volume/bottom water temperatures in the Pliocene (compared to the Pleistocene), orbital forcing remained an important driver of regional surface temperature variability, as well as seasonality changes, throughout the Pliocene. Not every warm interval within the Pliocene has equal relevance in understanding future climate change, since specific warm intervals (“interglacials”) occurred under conditions of orbital forcing unlike today. Orbital forcing is not a relevant forcing in Pliocene considerations of Climate Sensitivity, and should not be included in considerations of longer term climate sensitivity (Earth System Sensitivity; Willett et al., 2013; Haywood et al. 2013a; Martínez-Botí et al., 2015).

Climate modelling is enabling the Pliocene community to consider, in a more robust way, new targets for data acquisition and reconstruction, at the same time narrowing the relevant time window. This represents a significant shift away from deep time (pre-Quaternary) global geological reconstruction techniques and toward a methodological approach that is easily recognizable in Quaternary Science. In the near future, models and data will be more ‘*time consistent’*; meaning that models can be given more appropriate forcings to better represent the narrow time interval being reconstructed by geologists. Even so, limitations in correlation will continue. Models are helping to identify windows of time within the Pliocene (such as the interval immediately prior to and including Marine Isotope Stage KM5c) where orbital forcing was near modern and relatively stable (Haywood et al., 2013b). This means that orbitally-induced surface temperature variability will have a reduced potential to bias local temperature reconstructions, even if the correlation of the data to the peak of the interglacial event is in question (in this case KM5c; Haywood et al., 2013b).

The recent climate modelling study by Prescott et al. (2014) underlines the importance for enhanced chronological control and time consistent geological reconstructions to facilitate robust data-model comparison. Climate variability around two interglacial events within the mPWP that have different characteristics of orbital forcing were considered (Marine Isotope Stages KM5c and K1). Model results indicated that ±20 kyr on either side of MIS KM5c, orbital forcing produced a less than 1 °C change in global mean annual temperature. Regionally, mean annual surface air temperature (SAT) variability reached 2 to 3 °C. Seasonal variations exceed those predicted for the annual mean and locally exceeded 5 °C. Simulations ±20 kyr on either side of MIS K1, however, showed substantially increased variability in relation to KM5c due to the difference in orbital forcing at that time (Prescott et al., 2014).

Furthermore, Prescott et al. (2014) demonstrated that orbitally-forced changes in surface air temperature during interglacial events within the mPWP can be substantial and that uncertainties in correlation could easily contribute to patterns of apparent data/model disagreement. This is especially likely if geological archives (in the higher latitudes, for example) preserve a growing season signal rather than mean annual temperature. Climate model results indicated that peak MIS KM5c and K1 interglacial temperatures were not globally synchronous, highlighting leads and lags in temperature in different regions. This illustrates the pitfalls in aligning peaks in proxy temperatures across geographically diverse sites and indicates that a single climate model simulation for an interglacial event within the Pliocene is not adequate to capture peak temperature change in all regions (**Figure 6**).

1. **Summary** (311 words)

Based upon lessons learned from a long and productive relationship with the modelling community, Pliocene palaeoclimate reconstruction has undergone a paradigm change (Dowsett et al., 2013). The sophistication of fully-coupled climate models with high spatial resolution all but eliminates the need for global SST reconstructions with inherent shortcomings due to an insufficient number of localities available for contouring. SST studies are moving away from a global synthesis approach, focusing instead on locality-based regional reconstructions with emphasis on dynamics and processes. While multiple-proxy reconstructions of temperature are already standard, integration of these and other environmental proxies, whenever possible, will create a more complete and nuanced picture of the environment. This is being done by fully exploring the differences between proxies of the same environmental variable. For example in the marine realm, we have the ability to generate a warm season temperature and salinity, at the surface and at various depths, and fit those estimates within a framework of maximum and minimum annual temperature. These estimates can be complemented by measures of primary productivity and augmented by non-traditional signal carriers like Cheilostome Bryozoa providing a record of seasonality (Knowles et al., 2010). Sclerochronological analysis of selected mollusks calibrated with δ18O thermometry can provide a measure of seasonality and be a cross-check on paleontological and geochemical estimates of variability. This can be accomplished through high-resolution time series punctuated by environmental snapshots or through time-slices selected with an understanding of the need to minimize uncertainty in both experimental design and palaeoenvironmental reconstruction.

In the future, the utility of Pliocene data/model comparisons will also be greatly enhanced by 1) establishing a more precise chronology of the proxy data, 2) providing climate models with boundary conditions that are more consistent with the intervals of time studied for environmental reconstruction and 3) producing climate ensembles that more adequately capture orbital variability around any studied interval of the Pliocene.

**References**

Badger, M.P.S., Schmidt, D.N., Mackensen, A. and Pancost, R.D. High resolution alkenone palaeobarometry indicates relatively stable pCO2 during the Pliocene (3.3 to 2.8Ma). *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **371**: 20130094 (2013).

Ballantyne, A.P. et al. Significantly warmer Arctic surface temperatures during the Pliocene indicated by multiple independent proxies. *Geology* **38**, 603-606, doi:10.1130/G30815.1. (2010).

Barker, S. and Elderfield, H. Foraminiferal Calcification Response to Glacial-Interglacial Changes in Atmospheric CO2. *Science* **297**, 833-836 (2002).

Barron, J. A. in Pacific Neogene: Environment, Evolution, and Events (eds. R. Tsuchi and J. C. Jr. Ingle) 25-41, University of Tokyo Press (1992).

Barron, J. A. Pliocene paleoclimatic interpretation of DSDP Site 580 (NW Pacific) using diatoms. *Marine Micropaleontology* **20**, 23-44, doi:10.1016/0377-8398(92)90007-7 (1992).

Bartoli, G., Honisch, B. r. and Zeebe, R. E. Atmospheric CO2 decline during the Pliocene intensification of Northern Hemisphere glaciations. *Paleoceanography* **26**, PA4213, doi:10.1029/2010pa002055 (2011).

Berggren, W. A., Kent, D. V. and Van Couvering, J. A. in *The Chronology of the Geological Record* Vol. 10 (ed N. J. Snelling) 211-260 (Geological Society of London Memoir, 1985).

Bonham, S. G., Haywood, A. M., Lunt, D. J., Collins, M. and Salzmann, U. El Niño-Southern Oscillation, Pliocene climate and equifinality. *Philosophical Transactions of the Royal Society, A* **367**, 127-156, doi:10.1098/rsta.2008.0212 (2009).

Bragg, F. J., Lunt, D. J. and Haywood, A. M. Mid-Pliocene climate modelled using the UK Hadley Centre Model: PlioMIP Experiments 1 and 2. *Geoscientific Model Development* **5**, 1109-1125, doi:10.5194/gmd-5-1109-2012 (2012).

Brigham-Grette, J. and Carter, L. D. Pliocene marine transgressions of northern Alaska : circumarctic correlations and paleoclimatic interpretations. *Arctic* **55**, 74-89 (1992).

Brigham-Grette, J. *et al.* Pliocene Warmth, Polar Amplification, and Stepped Pleistocene Cooling Recorded in NE Arctic Russia. *Science* **340**, 1421-1427 (2013).

Chandler, M., Rind, D. and Thompson, R. Joint investigations of the middle Pliocene climate II: GISS GCM Northern Hemisphere results. *Global and Planetary Change* **9**, 197-219. (1994).

CLIMAP. The Surface of the Ice-Age Earth. *Science* **191**, 1131-1137, doi:10.1126/science.191.4232.1131 (1976).

Coates, A. et al. Closure of the Isthmus of Panama: The near-shore marine record of Costa Rica and western Panama. *Geological Society of America Bulletin* **104**, 814-828, doi:10.1130/0016-7606(1992)104<0814:cotiop>2.3.co;2 (1992).

Conte, M. H., Thompson, A. and Eglinton, G. Primary production of lipid biomarker compounds by Emiliania huxleyi: results from an experimental mesocosm study in fjords of southern Norway. *Sarsia* **79**, 319-332 (1994).

Cook C.P. et al. Sea surface temperature control on the distribution of far-traveled Southern Ocean ice-rafted detritus during the Pliocene. *Paleoceanography* **29**, 533-548. doi: 10.1002/2014PA002625 (2014).

Cronin, T. M. et al. Microfaunal Evidence for Elevated Pliocene Temperatures in the Arctic Ocean. *Paleoceanography* **8**, 161-173, doi:10.1029/93PA00060 (1993).

Crowley, T.J. Modeling Pliocene warmth. *Quaternary Science Reviews* **10**, 272-282. (1991)

Crowley, T.J. Pliocene climates: the nature of the problem. *Marine Micropaleontology* **27**, (1/4), 3-12. (1996).

deMenocal, P. B. African climate change and faunal evolution during the Pliocene-Pleistocene. *Earth and Planetary Science Letters* **220**, 3-24, doi:10.1016/S0012-821X(04)00003-2 (2004).

Dolan, A.M., et al. Sensitivity of Pliocene ice sheets to orbital forcing, *Palaeogeography, Palaeoclimatology, Palaeoecology* **309**, 98-110. doi: 10.1016/j.palaeo.2011.03.030, (2011).

Dolan, A.M. et al. Using results from the PlioMIP ensemble to investigate the Greenland Ice Sheet during the warm Pliocene. *Climate of the Past Discussions* **10**, 3483-3535 (2014).

Dowsett, H. J. and Cronin, T. M. High eustatic sea level during the middle Pliocene: Evidence from the southeastern U.S. Atlantic Coastal Plain. *Geology* **18**, 435-438, doi:10.1130/0091-7613(1990)018<0435:hesldt>2.3.co;2 (1990).

Dowsett, H. J. and Poore, R. Z. A new planktic foraminifer transfer function for estimating pliocene—Holocene paleoceanographic conditions in the North Atlantic. *Marine Micropaleontology* **16**, 1-23 (1990a).

Dowsett, H.J. and Poore, R.Z. Pliocene sea surface temperatures of the North Atlantic Ocean at 3.0 Ma. *Quaternary Science Reviews* **10**, 189-204. (1990b).

Dowsett, H. J. et al. Micropaleontological evidence for increased meridional heat transport in the North Atlantic Ocean during the Pliocene. *Science* **258**, 1133-1135, doi:10.1126/science.258.5085.1133 (1992).

Dowsett, H. J. et al. Joint investigations of the Middle Pliocene climate I: PRISM paleoenvironmental reconstructions. *Global and Planetary Change* **9**, 169-195, doi:10.1016/0921-8181(94)90015-9 (1994).

Dowsett, H., Barron, J. and Poore, R. Middle Pliocene sea surface temperatures: a global reconstruction. *Marine Micropaleontology* **27**, 13-25 (1996).

Dowsett, H. J. *et al.* Middle Pliocene Paleoenvironmental Reconstruction: PRISM 2. *U.S. Geological Survey, Open File Report* **99-535** (1999).

Dowsett, H. J. in Deep-time perspectives on climate change: marrying the signal from computer models and biological proxies (eds. M. Williams, A. M. Haywood, J. Gregory, and D. N. Schmidt) 459-480 (Micropalaeontological Society (Special Publication)), Geological Society of London (2007a).

Dowsett, H. J. Faunal re-evaluation of Mid-Pliocene conditions in the western equatorial Pacific. *Micropaleontology* **53**, 447-456, doi:10.2113/gsmicropal.53.6.447 (2007b).

Dowsett, H. J., Robinson, M. M. and Foley, K. M. Pliocene three-dimensional global ocean temperature reconstruction. *Climate of the Past* **5**, 769-783, doi:10.5194/cp-5-769-2009 (2009).

Dowsett, H. J. and Robinson, M. M. Mid-Pliocene equatorial Pacific sea surface temperature reconstruction: a multi-proxy perspective. *Philosophical Transactions of the Royal Society, A* **367**, 109-125, doi:10.1098/rsta.2008.0206 (2009).

Dowsett, H. et al. The PRISM3D paleoenvironmental reconstruction. *Stratigraphy* **7**, 123-139 (2010).

Dowsett, H. J. *et al.* Sea surface temperatures of the mid-Piacenzian Warm Period: A comparison of PRISM3 and HadCM3. *Palaeogeography, Palaeoclimatology, Palaeoecology* **309**, 83-91, doi:doi: 10.1016/j.palaeo.2011.03.016 (2011).

Dowsett, H. J. et al. Assessing confidence in Pliocene sea surface temperatures to evaluate predictive models. *Nature Climate Change* **2**, 365-371 (2012).

Dowsett, H. J. et al. Sea surface temperature of the mid-Piacenzian ocean: A data-model comparison. *Scientific Reports* **3**, 1-8, doi:10.1038/srep02013 (2013a).

Dowsett, H. J. et al. The PRISM (Pliocene Palaeoclimate) reconstruction: Time for a paradigm shift. *Philosophical Transactions of the Royal Society* **371**, 20120524. doi:10.1098/rsta.2012.0524 (2013b).

Dwyer, G. S. and Chandler, M. A. Mid-Pliocene sea level and continental ice volume based on coupled benthic Mg/Ca palaeotemperatures and oxygen isotopes. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **367**, 157-168, doi:10.1098/rsta.2008.0222 (2009).

Emiliani, C. Pleistocene temperature. *Journal of Geology* **63**, 538–578 (1955).

Fedorov, A. et al*.* Patterns and mechanisms of early Pliocene warmth. *Nature* **496**, 43-49 (2013).

Haug, G., Tiedemann, R., Zahn, R. and Ravelo, A. Role of Panama uplift on oceanic freshwater balance. *Geology* **29**, 207-210, doi:10.1130/00917613(2001) 029<0207:ROPUOO>2.0.CO;2 (2001).

Haywood, A.M., Valdes, P.J., Sellwood, B.W. Global scale palaeoclimate reconstruction of the middle Pliocene climate using the UKMO GCM: initial results. *Global and Planetary Change* **25**, (1/3), 239-256. (2000a).

Haywood, A.M., Sellwood, B.W., Valdes, P.J. Regional warming: Pliocene (3 Ma) paleoclimate of Europe and the Mediterranean. *Geology*, **28**, 1063-1066. (2000b).

Haywood, A. M., Valdes, P. J. and Sellwood, B. W. Magnitude of climate variability during middle Pliocene warmth: a palaeoclimate modelling study. *Palaeogeography, Palaeoclimatology, Palaeoecology* **188**, 1-24, doi:10.1016/S0031-0182(02)00506-0 (2002).

Haywood, A.M., Valdes, P.J. Modelling Pliocene warmth: contribution of atmosphere, oceans and cryosphere. *Earth and Planetary Science Letters* **218**, 363-377. doi: 10.1016/S0012-821X(03)00685-X, (2004).

Haywood, A.M., Valdes, P.J. Vegetation cover in a warmer world simulated using a dynamic global vegetation model for the Mid-Pliocene, *Palaeogeography, Palaeoclimatology, Palaeoecology* **237**, 412-427. doi: 10.1016/j.palaeo.2005.12.012, (2006).

Haywood, A.M., Dekens, P., Ravelo, A.C., Williams, M. Warmer tropics during the mid-Pliocene? Evidence from alkenone paleothermometry and a fully coupled ocean-atmosphere GCM, *Geochemistry, Geophysics, Geosystems*, **6**, Q03010. doi: 10.1029/2004GC000799, (2005).

Haywood A.M. et al. Are there pre-Quaternary geological analogues for a future greenhouse gas-induced global warming? *Philosophical Transactions of the Royal Society of London, Series A* **369**, 933-956. doi: 10.1098/rsta.2010.0317 (2011).

Haywood, A.M., et al. Large-scale features of Pliocene climate: Results from the Pliocene Model Intercomparison Project. *Climate of the Past* **9**, 191-209. doi: 10.5194/cp-9-191-2013, (2013a).

Haywood, A.M., et al. On the identification of a Pliocene time slice for data-model comparison. *Philos Trans A Math Phys Eng Sci,* **371**, pp.20120515. doi: 10.1098/rsta.2012.0515, (2013b).

Herbert, T. D., Peterson, L. C., Lawrence, K. T. and Liu, Z. Tropical Ocean Temperatures Over the Past 3.5 Million Years. *Science* **328**, 1530-1534, doi:10.1126/science.1185435 (2010).

Herbert, T., Peterson, L. and Gideon, N. Evolution of Mediterranean sea surface temperatures 3.5-1.5 Ma: regional and hemispheric influences. *Earth and Planetary Science Letters* (2015).

Hill, D. J., Haywood, A. M., Hindmarsh, R. C. A. and Valdes, P. J. in Deep time perspectives on climate change: marrying the signals from computer models and biological proxies (eds M. Williams, A. M. Haywood, D. Gregory, and D. N. Schmidt) 517-538 (*Micropalaeontological Society, Special Publication, Geological Society of London)*, (2007).

Hill, D. J., Csank, A. Z., Dolan, A. M. and Lunt, D. J. Pliocene climate variability: Northern Annular Mode in models and tree-ring data. *Palaeogeography, Palaeoclimatology, Palaeoecology* **309**, 118-127, doi:doi: 10.1016/j.palaeo.2011.04.003 (2011).

Howell, F.W., et al. Can uncertainties in sea ice albedo reconcile patterns of data-model discord for the Pliocene and 20th/21st centuries?, *Geophysical Research Letters*, **41**, 2011-2018. doi: 10.1002/2013GL058872 (2014).

Imbrie, J. and Kipp, N. G. A new micropaleontological method for quantitative paleoclimatology: application to a late Pleistocene Caribbean core. *The late Cenozoic glacial ages* **3**, 71-181 (1971).

IPCC. Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA, 1535 pp. (2013).

Ikeya, N. and Cronin, T. M. Quantitative Analysis of Ostracoda and Water Masses around Japan: Application to Pliocene and Pleistocene Paleoceanography. *Micropaleontology* **39**, 263-281 (1993).

Kennett, J. P. and Hodell, D. A. Stability or instability of Antarctic ice sheets during warm climates of the Pliocene? *GSA Today* **5**, 10-13 (1995).

Knowles, T. et al. Interpreting seawater temperature range using oxygen isotopes and zooid size variation in Pentapora foliacea (Bryozoa). *Marine Biology* **157**, 1171-1180, doi:10.1007/s00227-010-1397-5 (2010).

Kürschner, W. M., van der Burgh, J., Visscher, H. and Dilcher, D. L. Oak leaves as biosensors of late neogene and early pleistocene paleoatmospheric CO2 concentrations. *Marine Micropaleontology* **27**, 299-312, doi:10.1016/0377-8398(95)00067-4 (1996).

Laskar, J. *et al.* A long-term numerical solution for the insolation quantities of the Earth. *Astronomy and Astrophysics* **428**, 261-285 (2004).

Lawrence, K. T., Sosdian, S., White, H. E. and Rosenthal, Y. North Atlantic climate evolution through the Plio-Pleistocene climate transitions. *Earth and Planetary Science Letters* **300**, 329-342, doi:10.1016/j.epsl.2010.10.013 (2010).

Lea, D. W., Mashiotta, T. A. and Spero, H. J. Controls on magnesium and strontium uptake in planktonic foraminifera determined by live culturing. *Geochimica et Cosmochimica Acta* **63**, 2369-2379, doi:10.1016/S0016-7037(99)00197-0 (1999).

Lisiecki, L. E. and Raymo, M. E. A Pliocene-Pleistocene stack of 57 globally distributed benthic δ18O records. *Paleoceanography* **20**, doi:10.1029/2004PA001071 (2005).

Lunt, D. J., Valdes, P. J., Haywood, A. and Rutt, I. C. Closure of the Panama Seaway during the Pliocene: implications for climate and Northern Hemisphere glaciation. *Climate Dynamics* **30**, 1-18 (2008).

Lunt D.J., et al. Earth system sensitivity inferred from Pliocene modelling and data. *Nature Geosciences* **3**, 60-64. doi: 10.1038/NGEO706 (2010).

Lunt, D.J., et al. On the causes of mid-Pliocene warmth and polar amplification, *Earth and Planetary Science Letters* **321-322**, 128-138. (2012).

Lyell, C. Travels in North America; with geological observations on the United States, Canada, and Nova Scotia. J. Murray: London 2 vol. (J. Murray, 1845).

Martínez-Botí, M.A. et al. Plio-Pleistocene climate sensitivity evaluated using high-resolution CO2 records. *Nature* **518**, 49–54 (2015).

Masson-Delmotte, V., M. et al. Information from Paleoclimate Archives. In: *Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change* [Stocker, T.F., D. Qin, G.-K. Plattner, M. Tignor, S.K. Allen, J. Boschung, A. Nauels, Y. Xia, V. Bex and P.M. Midgley (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA. (2013).

Miller, K. G. *et al.* High tide of the warm Pliocene: Implications of global sea level for Antarctic deglaciation. *Geology* **40**, 407-410 (2012).

Miller, K. G. *et al.* The Phanerozoic Record of Global Sea-Level Change. *Science* **310**, 1293-1298, doi:10.1126/science.1116412 (2005).

Müller, P. J., Kirst, G., Ruhland, G., von Storch, I. and Rosell-Melé, A. Calibration of the alkenone paleotemperature index U< sub> 37</sub>< sup> K′</sup> based on core-tops from the eastern South Atlantic and the global ocean (60° N-60° S). *Geochimica et Cosmochimica Acta* **62**, 1757-1772 (1998).

Naish, T. R. and Wilson, G. S. Constraints on the amplitude of Mid-Pliocene (3.6–2.4 Ma) eustatic sea-level fluctuations from the New Zealand shallow-marine sediment record. *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences* **367**, 169-187 (2009).

O'Brien, C. L. *et al.* High sea surface temperatures in tropical warm pools during the Pliocene. *Nature Geoscience* **7**, 606–611, doi:DOI: 10.1038/NGEO2194 (2014).

Pagani, M., Liu, Z., LaRiviere, J. and Ravelo, A. C. High Earth-system climate sensitivity determined from Pliocene carbon dioxide concentrations. *Nature Geoscience* **3**, 27-30, doi:10.1038/ngeo724 (2010).

Pope, J.O., et al. Quantifying Uncertainty in Model Predictions for the Pliocene (Plio-QUMP): Initial Results, *Palaeogeography, Palaeoclimatology, Palaeoecology* **309**, 128-140. doi: 10.1016/j.palaeo.2011.05.004. (2011).

Prahl, F. G., Muehlhausen, L. A. and Zahnle, D. L. Further evaluation of long-chain alkenones as indicators of paleoceanographic conditions. *Geochimica et Cosmochimica Acta* **52**, 2303-2310, doi:10.1016/0016-7037(88)90132-9 (1988).

Prescott, C.L., et al. Assessing orbitally-forced interglacial climate variability during the mid-Pliocene Warm Period. *Earth and Planetary Science Letters* **400**, 261-271. (2014).

PRISM Project Members. PRISM 8°×10° Northern Hemisphere paleoclimate reconstruction: Digital data. *USGS Open File Report*, **94-281**, 23pp. (1994).

Raymo, M., Grant, B., Horowitz, M. and Rau, G. Mid-Pliocene warmth: stronger greenhouse and stronger conveyor. *Marine Micropaleontology* **27**, 313-326 (1996).

Rind, D. and Chandler, M. A. Increased ocean heat transports and warmer climate. *Journal of Geophysical Research* **96**, 7437-7461, doi:10.1029/91JD00009 (1991).

Rosenthal, Y., Lohmann, G. P., Lohmann, K. C. and Sherrell, R. M. Incorporation and Preservation of Mg in Globigerinoides sacculifer: Implications for Reconstructing the Temperature and 18O/16O of Seawater. *Paleoceanography* **15**, 135-145, doi:10.1029/1999PA000415 (2000).

Sachs, J. P., Pahnke, K., Smittenberg, R. and Zhang, Z. Biomarker Indicators of Past Climate. *Encyclopedia of Quaternary Science* (2013).

Salzmann, U., Haywood, A. M., Lunt, D. J., Valdes, P. J. and Hill, D. J. A new global biome reconstruction and data-model comparison for the Middle Pliocene. *Global Ecology and Biogeography* **17**, 432-447, doi:10.1111/j.1466-8238.2008.00381.x (2008).

Salzmann, U. et al. Challenges in Reconstructing Terrestrial Warming of the Pliocene Revealed by Data-Model Discord, *Nature Climate Change* **3**, 969-974. doi: 10.1038/nclimate2008. (2013).

Seki, O. et al. Alkenone and boron-based Pliocene pCO2 records. *Earth and Planetary Science Letters* **292**, 201-211, doi:10.1016/j.epsl.2010.01.037 (2010).

Shackleton, N. J., Hall, M. A. and Pate, D. Pliocene stable isotope stratigraphy of Site 846. *Proceedings of the Ocean Drilling Program, Scientific Results* **138**, 337-355, doi:10.2973/odp.proc.sr.138.117.1995 (1995).

Sloan, L. C., Crowley, T. J. and Pollard, D. Modeling of middle Pliocene climate with the NCAR GENESIS general circulation model. *Marine Micropaleontology* **27**, 51-61, doi:10.1016/0377-8398(95)00063-1 (1996).

Sohl, L. E. et al. PRISM3/GISS topographic reconstruction. *Geological Survey Data Series* **419**, 6p. (2009).

Sosdian, S. and Rosenthal, Y. Deep-Sea Temperature and Ice Volume Changes Across the Pliocene-Pleistocene Climate Transitions. *Science* **325**, 306-310 (2009).

Thompson, R.S. and Fleming, R.F. Middle Pliocene vegetation: reconstructions, paleoclimatic inferences, and boundary conditions for climatic modeling. *Marine Micropaleontology* **27**, 13-26. (1996).

Tripati, A. K., Roberts, C. D. and Eagle, R. A. Coupling of CO2 and ice sheet stability over major climate transitions of the last 20 million years. *Science* **326**, 1394-1397, doi:10.1126/science.1178296 (2009).

Unger, N., and X. Yue. Strong chemistry-climate feedback in the Pliocene. *Geophysical Research Letters* **41**, 527–533 (2014).

Valdes, P. Built for stability. *Nature Geoscience* **4**, 414-416, doi:doi:10.1038/ngeo1200 (2011).

Vaughan, A.P.M. Climate and geology – a Phanerozoic perspective. In *Deep time perspectives on climate change: marrying the signals from computer models and biological proxies* (eds M. Williams, A. M. Haywood, D. Gregory, and D. N. Schmidt) 517-538 Micropalaeontological Society, Special Publication, Geological Society of London 5-59 (2007).

Volkman, J. K., Eglinton, G., Corner, E. D. and Forsberg, T. Long-chain alkenes and alkenones in the marine coccolithophorid Emiliania huxleyi. *Phytochemistry* **19**, 2619-2622, doi:10.1016/S0031-9422(00)83930-8 (1980).

Wara, M. W., Ravelo, A. C. and Delaney, M. L. Permanent El Niño-Like Conditions During the Pliocene Warm Period. *Science* **309**, 758-761, doi:10.1126/science.1112596 (2005).

Wardlaw, B. R. and Quinn, T. M. The record of Pliocene sea-level change at Enewetak Atoll. *Quaternary Science Reviews* **10**, 247-258, doi:10.1016/0277-3791(91)90023-N (1991).

Willeit M, Ganopolski A, Feulner G. On the effect of orbital forcing on mid-Pliocene climate, vegetation and ice sheets. *Climate of the Past* **9**, 1749–1759. doi:10.5194/cp-9-1749-2013 (2013).

Zhang, Y.G., Pagani, M. and Liu, Z. A 12-Million-Year Temperature History of the Tropical Pacific Ocean. *Science* **344**, 84-87. (2014).

Zhang Z-S. et al. Mid-pliocene Atlantic Meridional Overturning Circulation not unlike modern, *Climate of the Past* **9**, pp.1495-1504. doi: 10.5194/cp-9-1495-2013. (2013).

Zubakov, V. A. and Borzenkova, I. I. Pliocene palaeoclimates: Past climates as possible analogues of mid-twenty-first century climate. *Palaeogeography, Palaeoclimatology, Palaeoecology* **65**, 35-49, doi:10.1016/0031-0182(88)90110-1 (1988).

**Acknowledgments**

AH acknowledges that this review was completed in receipt of funding from the European Research Council under the European Union’s Seventh Framework Programme (FP7/2007-2013)/ERC grant agreement no. 278636. HD acknowledges the US Geological Survey Climate and Land Use Change Research and Development Program and the John Wesley Powell Center for Analysis and Synthesis. AH and HD thank Aisling Dolan, Caroline Prescott, Marci Robinson and Lynn Wingard for their comments on a previous version of the review.

**Box1. Pliocene Research Interpretations and Synoptic Mapping** (477 words)

The initial PRISM reconstruction was confined to the North Atlantic region. Palaeotemperature estimates were based upon quantitative analysis of foraminifer and ostracode faunas from deep-sea cores and continental sections. Even this early and relatively primitive analysis documented a level of polar amplification of surface temperatures unseen today (Dowsett et al., 1992). With the addition of marine data from the North Pacific (e.g., Barron, 1992a/b; Ikeya and Cronin, 1993; Cronin et al., 1993) and for the first time vegetation or land surface cover (PRISM, 1994), sea level and ice volume/distribution estimates, PRISM was extended to a Northern Hemisphere Reconstruction (Dowsett et al., 1994: PRISM, 1994). Since that time each successive phase of the project (PRISM1 through 3) has added a new aspect of the palaeoenvironment or refined existing data sets (Dowsett et al., 1996, 1999, 2010; Thompson and Fleming, 1996).

The chronostratigraphic framework within which a palaeoenvironmental reconstruction is developed is fundamental to its uncertainty. When the PRISM project commenced, confident correlation of sequences from one ocean basin to another was limited by the resolution of the palaeomagnetic time scale and temporally calibrated fossil evolutionary first and last appearance events. This limitation and initial observations from the North Atlantic reconstruction defined a 300,000 year time slab within which long distance correlations were possible, and the warm phases of climate could be averaged. As the magnitude of the Pliocene temperature anomaly became better defined, new methods of standardization (e.g. calibration of various palaeothermometres to the same modern temperature dataset) and refinement of the chronologic limits of the time-slab using newly developed oxygen isotope stratigraphy (Lisiecki and Raymo, 2005), PRISM consistently improved the overall confidence of the reconstruction.

The PRISM3D global reconstruction consists of six discreet data sets representing different aspects of the Pliocene world: sea surface temperatures, deep ocean temperatures, land vegetation, ice distribution, topography and sea level (Dowsett, 2007; Dowsett et al., 2009a/b, 2010; Hill et al., 2007; Salzmann et al., 2008; Sohl et al., 2009). It differs from its predecessors in that it includes a subsurface ocean temperature reconstruction, integrated geochemical sea surface temperature proxies to supplement faunal-based estimates, and used numerical models for the first time to augment palaeobotanical data in the creation of a biome-based land cover scheme.

The PRISM4 reconstruction has new topography and ocean bathymetry, ice sheets, sea level, soils and lakes data sets (Dowsett et al., in prep.). SST and BIOME land cover remain the same as in PRISM3; new high-resolution time series are being developed to provide information on variability. These elements of the PRISM4 reconstruction will be best suited for evaluation of new climate model simulations that will be carried out as part of the second phase of PlioMIP (PlioMIP2).

**Figure captions**

**Figure 1**. The PRISM interval in relation to the long-term climate evolution of the Late Pliocene as shown in the LR04 benthic oxygen isotope stack of Lisiecki and Raymo (2005). Vertical dashed line shows present day δ18O value. The PRISM3 warm interval (3.264 to 3.025 Ma) is shown by the horizontal shaded grey bar. Blow up on right shows details of LR04 and position of PRISM4/PlioMIP2 focus. Obliquity (°), eccentricity, precession and July insolation at 65°N (W m-2) are also shown for reference (Laskar et al., 2004).

**Figure 2**. Time series illustrating commonly used palaeoenvironmental proxies. A. Equatorial Atlantic SST based on alkenone unsaturation index (Herbert et al., 2010); B. Carribean Sea oxygen isotope record of Haug and Tiedeman (2001) showing increasing salinity due to shoaling of the CAS; C. Terrestrial record of mean warmest month temperature from Lake El’gygytgyn (Briggham-Grette et al., 2013); D. Equatorial Pacific SST records based upon Mg/Ca paleothermometry (Wara et al., 2005); E. Estimates of Pliocene atmospheric CO2, dark blue dots, δ13C, Raymo et al. (1996); blue, Ba/Ca, Tripati et al. (2009); red, alkenone, Pagani et al. (2010); orange, stomata, Kürschner et al. (1996); green, δ11B, Bartoli et al. (2011); yellow, δ11B, Seki et al. (2010); pink, alkenone, Seki et al. (2010); Modified from Fedorov et al. (2013).

**Figure 3**. PRISM mean annual SST anomalies. (A) PRISM0 Northern Hemisphere reconstruction (Dowsett et al. 1994); (B) PRISM2 global reconstruction (Dowsett et al., 1999; Dowsett, 2007a); (C) PRISM3 confidence assessed global reconstruction where larger diameter circles represent higher confidence levels (Dowsett et al., 2012; 2013).

**Figure 4**. Pliocene simulations. (A) GISS AGCM with PRISM0 prescribed boundary conditions (Chandler et al., 1994); (B) HadCM3 AOGCM initiated with PRISM2 boundary conditions (Haywood and Valdes, 2004); (C) HadCM3 AOGCM PlioMIP simulation initiated with PRISM3D boundary conditions (Bragg et al., 2012).

**Figure 5**. IPCC Pliocene data/model comparison modified from Masson-Delmotte et al. (2013). Comparison of PRISM3 proxy data and the PlioMIP multi-model mean (MMM) simulation, for (far left) sea surface temperature (SST) anomalies, (middle left) zonally averaged SST anomalies, (middle right) zonally averaged global (green) and land (grey) surface air temperature (SAT) anomalies and (right) land SAT anomalies. Zonal MMM gradients (shown in middle left and middle right) are plotted with a shaded band indicating 2 standard deviations. Site specific temperature anomalies estimated from PRISM proxy data are calculated relative to present site temperatures and are plotted (in left and right) using the same colour scale as the model data, and a circle-size scaled to estimates of data confidence (data from Haywood et al., 2013a; Dowsett et al., 2012 and Salzmann et al., 2013).

**Figure 6**. HadCM3 data exploring climate variability during the Pliocene redrawn from Prescott et al. (2014). Annual mean Pliocene Surface Air Temperature (SAT ºC) predictions for (A) interglacial Marine Isotope Stage (MIS) KM5c minus a pre-industrial experiment and (B) interglacial MIS K1 minus a pre-industrial experiment. Note the difference in temperature response between the two interglacial events with MIS K1 having a larger response than MIS KM5c. Also shown is the maximum annual SAT change derived from ten orbital sensitivity simulations differenced from the MIS KM5c (C) and the MIS K1 (D) control simulations in each grid square. Note the larger degree of variability 20 Kyrs ± of the peak of MIS K1 compared to 20 Kyrs ± of the peak of KM5c. Maximum SAT shown around interglacial peak MIS KM5c (E) and K1 (F). Colours denote model simulations in which maximum temperature occurred per model grid square. This indicates that maximum temperature for each interglacial was not synchronous and also varied between MIS KM5c and K1.

**END.**