Response to reviewers

Quantifying late-Holocene climate in the Ecuadorian Andes using a chironomid-based temperature inference model

We wish to thank everyone who contributed to the improvement of this manuscript, specifically the comprehensive suggestions from the two reviewers. Outlined below is a detailed response to both the reviewers' general, and specific comments. The manuscript has been substantially changed as a result of the reviewer's general comments (see below; *response to general comments*) and minor edits have been corrected and recorded accordingly. All typos and minor formatting errors, highlighted in the annotated pdf have also been changed.

Response to General comment: *reliability of the reconstruction*.

Whilst both reviewers, and an independent author who contributed to the online discussion, commented on the value of the study, all had major concerns relating to the final environmental reconstruction from Laguna Pindo and the subsequent interpretation. We would agree with all the reviewers that the development of a chironomid transfer function for the tropical Andes is an important contribution to tropical palaeolimnology and paleoclimatology. We also acknowledge, however, that chironomid studies from the tropics remain rare and little is known about the autecology of many of the taxa, namely their ecological tolerances relating to climatic variables. As a result, the environmental reconstruction from Laguna Pindo has some issues, namely unrealistically cold temperatures and significant inter sample variability. We would agree with all reviewers, and acknowledge in the original manuscript, that many of these fluctuations most likely relate to issues with the transfer function and/or the fossil record. These could include:

- The response to secondary variables, namely precipitation.
- Low head capsule concentrations in many of the samples.
- Un-even distribution of calibration lakes over the environmental gradient due to the steep topography of the Andes.
- Taxonomic issues (i.e different species between fossil samples and modern samples that currently cannot be separated using only larval head capsule material).

Many of these limitations are discussed in the manuscript. Indeed, we would argue a central point of this work would be the comparison of WA and Bayesian methods, in order to further explore these limitations. The application of the Bayesian model results in a less variable reconstruction, and an explanation for why the uncertainty associated with the reconstruction is greater than the climate variability we are reconstructing. The individual likelihood function of fossil taxa and the resulting posterior probability distribution for temperature sheds light on how the un-even distribution of calibration lakes, and subsequent skewed distribution of taxa, is affecting the inferred temperatures. Some taxa may have under-estimated temperature optima due to the over representation of cold lakes in the calibration datasets and this may have had a significant effect on the temperature reconstruction, in particular the under-estimated Holocene temperatures. The error associated with both reconstructions is entirely consistent with a constant temperature of 20°C. We would agree with the reviewers that attributing the variability of the reconstruction to anything more than noise would be an overstatement at this point. This work does, however, suggest the way forward for improving temperature reconstructions, namely, improving the richness/sampling of the training set to enable the detection of smaller signals.

The anonymous reviewer provides some options for rectifying these issues:

> *"(i) The TF is optimized, undergoes additional testing, the quality of the reconstruction is substantially improved (robustness of the TT amplitudes, robustness of the cold anomalies, etc.) and/or (ii) the profile of the reconstruction is lowered; given the pertinent deficits the reconstruction is qualitative and NOT quantitative, not overstating the results and conclusions."*

Unfortunately, we do not feel that the quality of the reconstruction can be substantially improved at this stage for many of the reason discussed previously and therefore we cannot meet the first criteria (i). Meeting this criterion would involve extensive fieldwork in an attempt to source more calibration sites along the temperature gradient. For this reason we propose to move forward with the reviewers second suggestion, i.e. lowering the profile of the reconstruction in the manuscript. We feel this option will allow the manuscript to make a meaningful contribution to the literature, whilst honestly representing the current sate of chironomid research in the area and addressing many of the concerns of all reviewers and online contributions.

The following major changes have been made to the manuscript in order to address the general comments of both reviewers:

i) The Introduction has been shortened (L.126-L132) in order to reflect the new focus of the manuscript, i.e. refining the proxy as opposed to palaeoclimate inferences. This modification also addresses a concern of the anonymous review, which noted the introduction as being overly long.

ii) Sub section 5.4 *Laguna Pindo temperature reconstructions* has been removed. This subsection is no longer needed as the temperature reconstruction is presented as qualitative and only used to further understand the various models.

iii) Sections 5.6 (*Cooling climate 3800-2800 cal yrs BP)*, and 5.7 (*Recent cooling)* have been removed. Based on the recommendations of the reviewers, we have changed the focus of the manuscript to center on proxy development not palaeoclimate interpretations. The Laguna Pindo reconstruction is used to understand the limitations of each model and is presented as a more qualitative interpretation of climate variability over this time period. The conclusions of the manuscript focus on our future recommendations for improving palaeotemperature inferences using chironomids. This addresses the current limitations of Neotropical palaeolimnology using chironomids, and provides a list of necessary criteria for future researchers wishing to explore this proxy further.

Response to specific comments

Line Number	Reviewers comment	Response
L. 38 R2	**General Comment** Due to the limitations of the environmental reconstruction the anonymous review suggest reference to the reconstruction should not be "quantitative"	*"…the first quantitative reconstruction…"* has been removed
L. 349 R1	I do not think that that lakes located between 1000 and 300 m asl can be considered to be similar in elevation to Laguna Pindo (~ 1200 m asl). Applying a standard lapse rate to this elevation range suggests that MAT for the lowest and highest lakes would vary by 12-20oC.	"*…although the samples plot within the range of modern calibration lakes that lie at similar elevations (1000-3000 m a.s.l).*" has been removed
L. 415 R1	Reporting the RMSEP as a % of the total MAT range captured by the training set would be useful.	Sentence becomes; *"Although both models (WA inverse and Bayesian) perform well (WA RMSEP= 2.4°C/ 9.6% of training set range and Bayesian RMSEP= 2.3°C/9.2% of training set range)…"*
Fig 1 R1	requires a N-arrow	N arrow has been added
Fig 2 R1	"PH" should be corrected.	PH has been corrected to pH
R1	It is not clear why non-limnological variables such as latitude were included in the exploratory analysis. Latitude, longitude and elevation are not directly controlling the distribution of midges; the analyses should be re-run with only environmental variables that have the potential to directly control the distribution of midges included.	These variables have been excluded from the analysis.
R2	The Introduction could be shortened (quite lengthy).	L.126-L132 have been removed from the introduction and the manuscript shortened to reflect the new direction of the paper.
R2	Chapter 3: I would not make too many sub-chapters (only one paragraph in 3.4 and 3.5)	Sub headings 3.4; 3.5; 3.6; and 3.7 have been removed
R2#	The sampling design for the downcore analysis should be described in detail (continuous, discrete sampling, regular intervals, stratigraphically. . .?) What is the percentage of sediment that is actually covered in the analysis? (e.g. 1 cm slice every 10 cm sediment makes 10% coverage and 90 % is not covered; this has serious implications regarding the robustness of the	Sentence added; *"The sampling interval for chironomid analysis was not uniform due to a varying sedimentation rate. To achieve as even a coverage possible over the time interval, samples were taken between every 10 and 20cm."*

	reconstruction).	
R2	I would include the Suppl Fig (Chronology) in the manuscript.	Table S1 and Figure S1 have know been included in the manuscript.
R2	Title: reconcile. It is quantitative indeed, but how robust and how good are the numbers? => Qualitative	Title has been changed to; Inferring late-Holocene climate in the Ecuadorian Andes using a chironomid-based temperature inference model
L.77 R2	Shulmeister	Corrected
L.93 R2	... preceding Glacial and Late-Glacial period ... (if you refer to 25-11.7 kBP; 20-25 kPB is not Late Glacial)	Changed to; *"(… c. 15,000-11,700 years before present…)"*
L.95 R2	Make also reference to Marcott et al 2013. This is the most comprehensive dataset.	Marcott et al 2013 has been added.
L.99-100 R2	Growing evidence from the tropics? I'm not sure about this. In fact it is still very controversial whether cold events (depending on the time scale) were globally, hemispherically or regionally synchronous (Wanner et al. 2011 QSR, Neukom et al 2014. NatCC;PAGES 2k 2013). The PAGES 2k Consortium 2013 has shown that, with a few exceptions (with solar-volcanic downturns) multi-decadal long cold phases were not coherent across the globe. Maybe rephrase sentence.	Changed to; "Some evidence from the tropics suggests Holocene climate fluctuations such as the LIA maybe global events…."
L.108-109 R2	References not appropriate (these are not climatology papers). Make ref- erence to Garreaud et al 2009 or Stefan Hastenrath 1991 Climate Dynamics of the Tropics or similar.	References have been modified.
L.151 R2	How reliable are WTs in a 10 cm deep water body? It should be assessed how sensitive the TF is with/without such lakes. In such water bodies the difference between MAT and WT is typically very large (in particular Tmax). I guess that the TF stats could be improved.	In producing the manuscript we ran the transfer function using multiple combinations of different lakes included and excluded. This included removing very shallow lakes and overly deep lakes. The results presented are for the best performing inference model. We believe the problems that are leading to the unreliable reconstruction are overwhelmingly those discussed with reference to the all reviewers' general comments. The manuscript has been changed accordingly to address this.
L.154 R2	. . . uppermost 1-2 cm . . . representing 5-20 years. . . Well, it was done like this and is	As pointed out by the reviewer the sampling method adopted here is

	usually done like this. But this implies that the sample for the Training Set depicts in one lake interannual/subdecadal variability (which may be very different from climatology!) and in another sample it is rather climatology (20-30 yrs). I suspect that this adds substantial errors to use the uppermost 3-4 cm of sediment to make sure that 20-30 yrs (climatology) are represented. The TT trends during 30 yrs are relatively small and similar in all lakes of the training set.	common practice for chironomid studies of this kind. We would agree with the reviewer that testing the results of various sampling methods would be a worthy endeavor. The reviewer makes an important point that the uppermost sediments likely reflect inter-annual variability whilst deeper homogenous sampling may more accurately reflect climatology. Addressing this directly, however, would call for a complete re-sampling of the entire calibration dataset and will very probably not address the central concern of the reviewers; improving the reconstruction. This suggestion would not reduce the problems associated with un-even sampling.
L.171 R2	Fourteen 14C samples? Fig SOM shows six of them. Where are the others? Pls change and make it consistent with L 324 ff.	The Laguna Pindo record is much older and longer than the portion presented here. Much of the record is radiocarbon infinite and work is on-going to produce a complete age depth model. Presented here is the portion of the record for which chironomid remains are found. This is addressed in Line 330-332 “*The best-fit age depth model for Laguna Pindo was a smooth spline (Fig S1). Due to the absence of chironomids at the bottom of the sequence, six radiocarbon samples were used for building the model with a total depth of the sediment considered of 461 cm”* . This figure has been removed from the SOM and placed in the manuscript itself.
L188 R2	Were nutrients (N and P) not measured? This might be a problem (Lotter et al 1998 J Paleolimnology)	Samples were taken for nutrients (anions and cation). Although filtered in the field, due to the remoteness of the fieldwork and continued biological activity, these samples were no longer reliable once returned for laboratory analysis in the UK.
L272 and 275 R2	avoid references in the results section. This reads like ‘Discussion’ L280 . . . optimum. . . (?)	References have been removed *“…optima”* changed to *“…optimum”*
L295 R2	in general, hc counts should be given in all Figures and Tables.	Hc counts has been added to all figures. The total number of head capsules for each calibration lake can be found in the data archive or *Matthews-bird et al 2015*
L295 f	is rather Discussion than Results. Move this paragraph.	Paragraph has been moved to

		discussion
L300 R2	Yes, this is critical (number of hc). It should be assessed whether the number of hc has an effect on the calibration statistics, in particular the residuals. (see also L303, I am not sure if this is the only criterion according to which the TF could be optimized)	We agree that the affect of head capsule concentration is extremely important.
L.302 R2	Table 1 does not show these details (which are important), Table 1 shows the summary only. The details (hc) should be given (in the SOM)	Total number of head capsules for each lake can be found in the data repository.
L.327 R2	The sampling design must be clarified (in the Methods section). You took 30 samples spread over 420 cm. How did you take the samples? 1 cm slice every 10-15 cm? Stratigraphically (according to which criteria?) or continuously (complete sediment section)?	This has been addressed by a previous comment R2#.
L.353 R2	.. only seven samples? According to Fig 9 and the vertical dashed line there are many more.	This was a typo that has been rectified, 14 samples have a poor fit to temperature.
L.485 R2	I don't think that anything is known about the precip/temperature relationship during the Late Holocene.	Sentence changed to; *"The location of Laguna Pindo makes it a good palaeoecological setting to record the response of temperature-sensitive proxies"*
L.495 R2	I think this is a substantial problem.	We agree with the reviewer that the lack of modern analogues is a substantial problem with the reconstruction. This is honestly discussed in the paragraph cited. This lack of modern analogues most likely reflects the uneven distribution of calibration lakes and the particular lack of lakes surrounding the fossil site.
L. 531 R2	according to this statement I would conclude that the temperature reconstruction of Laguna Pindo is qualitative at best.	We agree with the reviewer that more work is needed before Neotropical chironomids can be described as quantitative indicators of temperature. As highlighted in our response to the general comments we accept that the profile of the reconstruction should be lowered. The passage now reads; *"The WA inverse MAT reconstruction, however, is statistically significant based on the criteria described by Telford and Birks (2011a) (Fig 10) suggesting that despite precipitation as a possible confounding variable, a temperature signal can be*

		obtained from Neotropical chironomids although we caution against an over interpretation at this stage. Due to some of the limitations discussed previously, the reconstruction can currently only be deemed qualitative and requires more research."
L. 539 R2	maybe also refer to Kanner et al (speleothems) and Ledru et al (N Ecuador)	This section has now been removed and significantly modified. The manuscript no longer over interprets the final reconstruction and these suggestions are no longer relevant.
L548 R2	Jones & Mann 2004 is not the best (has been criticized; S-Hemisphere is very poor). Suggestion: PAGES 2k 2013.	See previous comment
L.552/553 R2	I don't think this is true. There's a large body of literature pointing out the role of volcanoes, or a combination of S+V . . . rephrase sentence.	See previous comments
L.555 R2	No, I don't think this is true (cool from 400 yr BP onwards). The sample at 250 yr BP is still among the warmest of the entire record, almost as warm as today (!). There is only 1 sample (at 1850 AD) that shows cool conditions, and it is very questionable how robust that is (see your comment and my comments above)	See previous comment
L.569ff R2	It has been repeatedly demonstrated that the Andean ice cores (stable iso- topes) record precipitation and not temperature (as claimed by Thompson et al).	See previous comment
	LIA. Yes, this value has been reported for two Venezuelan glaciers (at 4600 and 5000 masl, mainly inferred from a drop in ELA by 300-500 m; Polissar et al. 2006). I doubt that similar (special high-elevation) conditions apply for L Pindo, given the limitations of the reconstruction (see above). This value seems extraordinarily high to me. Alternatively an explanation should be provided showing that such large TT amplitudes are physically plausible at local scales.	See previous comment
L.598 R2	Yes, the potential is shown (with the TF). But the reconstruction has major problems and severe limitations (see above). I would say: qualitative at best.	We agree with the reviewer and have modified the manuscript accordingly
L.605 R2	: . . .). Special. . .	Rectified
L624 R2	reference listed twice L634: Dryas-Holocene L667: check carefully L668: . . .Science 289	Duplicate reference removed
L.680.	Vol missing	Volume added

R2		
L.702 R2	Lemke	Rectified
L.810 R2	... Science 234, ...	Rectified
L.814 R2	Ref listed twice (also L819)	Duplicate reference removed
L. 818 R2	Holocene	Rectified
L.841 R2	Woodward, C	Rectified
L.855 R2	LOI: specify 550 or 950; ditto L858, Table 1 and Table 2, L883	550 has been specified
Table 1	Data set should be made available in full detail	The research was funded by the Natural Environment Research Council UK (NERC). All data will be uploaded to the NERC data repository before publication
Table 3:	add units (where appropriate), also Caption Fig 5	Units added
Fig 2:	pH	PH changed to pH
Fig 3	(all Figs where appropriate, Fig 6, Fig 9): numbers of hc should be shown. It would be interesting to see the ‘unusual lakes’ (e.g. those with water depth of 10 cm).	Hc has been added to all necessary figures; fig 3,6,9

Inferring late-Holocene climate in the Ecuadorian Andes using a chironomid-based temperature inference model

frazer 2/26/2016 3:21 PM
Deleted: Quantifying

Frazer Matthews-Bird[1&2], matthewsbirdf@fit.edu

Stephen J. Brooks[3], S.Brooks@nhm.ac.uk

Philip B. Holden[1], philip.holden@open.ac.uk

Encarni Montoya[1], encarni.montoya@open.ac.uk

William D. Gosling[1, 4], W.D.Gosling@uva.nl

[1]Department of Earth, Environment & Ecosystems, The Open University, Walton Hall, Milton Keynes, MK7 6AA, UK.

[2] Biological Sciences, Florida Institute of Technology, 150 West University Boulevard, Melbourne, FL 32901, USA

[3]Department of Life Sciences, Natural History Museum, Cromwell Road, London SW7 5BD, UK.

[4]Palaeoecology & Landscape Ecology, Institute for Biodiversity & Ecosystem Dynamics, University of Amsterdam, P.O. Box 94248, 1090 GE Amsterdam, The Netherlands

Corresponding author: Frazer Matthews-Bird

Key words: Bayesian, weighted-averaging, transfer function, chironomids, Holocene climate change, Ecuador

Abstract

Presented here is the first chironomid calibration dataset for tropical South America. Surface sediments were collected from 59 lakes across Bolivia (15 lakes), Peru (32 lakes) and Ecuador (12 lakes) between 2004 and 2013 over an altitudinal gradient from 150 m above sea level (a.s.l) to 4655 m a.s.l, between 0-17°S and 64-78°W. The study sites cover a mean annual temperature (MAT) gradient of 25°C. In total, 55 chironomid taxa were identified in the 59 calibration data-set lakes. When used as a single explanatory variable, MAT explains 12.9% of the variance (λ_1/λ_2= 1.431). Two inference models were developed using weighted averaging and Bayesian methods. The best performing model using conventional statistical methods was a WA (inverse) model (R^2_{jack}= 0.890, $RMSEP_{jack}$ = 2.404, Mean $bias_{jack}$= -0.017, Max bias $_{jack}$=4.665). The Bayesian method produced a model with R^2_{jack}= 0.909, $RMSEP_{jack}$ =2.373, Mean $bias_{jack}$= 0.598, Max bias $_{jack}$= 3.158. Both models were used to infer past temperatures from a *c.* 3000 yr record from the tropical Andes of Ecuador, Laguna Pindo. Inferred temperatures fluctuated around modern day conditions but showed significant departures at certain intervals (*c*. 1600 cal yr BP; *c*. 3000-2500 cal yr BP). Both methods (WA/Bayesian) showed similar patterns of temperature variability; however, the magnitude of fluctuations differed. In general the WA method was more variable and often under-estimated Holocene temperatures (*c.* -7±2.5°C relative to the modern). The Bayesian method provided temperature anomaly estimates for cool periods that lay within the expected range of the Holocene (*c.* -3±3.4°C). The error associated with both reconstructions is consistent with a constant temperature of 20°C for the past 3000 yrs. Increasing the number, and spread, of lakes in the calibration dataset would enable the detection of smaller climate signals.

1. Introduction

Holocene climate variability (11.7 kcal yrs BP – present) offers the most recent opportunity to parameterise climate and ecosystem responses to natural forcing under current boundary conditions in the absence of intense anthropogenic activity (Mayewski *et al*., 2004; Oldfield and Steffen, 2014). Furthermore, quantitative estimates of past climate over long time scales (>1000 yrs) are vital to improving the reliability of modelling and prediction of present and future climate variability (Mayewski *et al*., 2004). The spatial distribution of palaeoclimate records, however, is currently uneven around the world. Quantitative reconstructions of past climate are common from mid- to high- latitudes of both hemispheres but data is much scarcer from low-latitude (tropical) regions (Jansen *et al.*, 2007). Tropical climate is the dominant driver of atmospheric circulation (Ivanochko *et al., 2005)* and the source of intermittent phenomena, such as the El Niño Southern Oscillation (ENSO), which has a global influence on climate (Collins *et al*., 2010). Quantitative estimates of past climate from the low latitude tropics, therefore, are crucial for investigating not only regional climate processes, but also teleconnections on long timescales (>1000 years) (Garreaud *et al.*, 2009; Jomelli *et al*., 2009; Vuille *et al*., 2000). Here we develop the first chironomid-based temperature inference model for tropical South America. The model is applied to a Holocene lake sediment sequence to generate a chironomid-inferred temperature reconstruction from the tropical East Andean flank.

frazer 2/25/2016 10:27 AM
Deleted: the first quantitative

frazer 2/25/2016 10:27 AM
Deleted: s

Chironomidae (non-biting midges) is a family of two-winged aquatic insects of the order Diptera. The family is globally distributed and one of the most diverse within aquatic ecosystems (Armitage *et al*., 1995). Many species are stenotopic, and their short life-cycles and ability to colonise favourable regions quickly means the insects are extremely sensitive to environmental change (Pinder, 1986). The head capsules of chironomid larvae are well preserved in lake sediments and have been used extensively as palaeoecological proxies (Brooks, 2006; Walker and Cwynar, 2006). Chironomid-based temperature inference models, derived from modern calibration data sets, have been applied across North America (reviewed in Walker and Cwynar, 2006), Eurasia (reviewed in Brooks, 2006), and more recently the method has been applied in the Southern Hemisphere in Patagonia (Massaferro and Larocque, 2013; Massaferro *et al*., 2014), Central America (Wu *et al*., 2014), East Africa (Eggermont *et al*., 2010), and Australasia (Dimitriadis and Cranston 2001; Woodward and Shulmeister 2006).

frazer 2/25/2016 1:38 PM
Deleted: r

Transfer functions make a number of underlying assumptions; particularly the environmental variable to be reconstructed is an ecologically important determinant in the system, and environmental variables other than one being reconstructed have a negligible effect on species assemblages (Juggins, 2013). Rarely are ecological systems as simple as transfer functions would imply and violations of these assumptions will undermine the validity of the environmental reconstruction (Juggins, 2013). Nevertheless, despite known inherent problems associated with transfer functions (Huntley, 2012; Juggins, 2013; Velle *et al*., 2010), quantitative reconstructions from chironomid assemblages often produce consistent results that compare well with other proxy estimates of past temperature (Brooks, 2000; Brooks *et al.*, 2012; Heiri *et al.*, 2007). The best performing inference models can reconstruct temperatures with errors of *c.* 1°C (Brooks and Birks, 2001; Eggermont et al., 2010; Heiri et al., 2003; Olander et al., 1999a; Rees et al., 2008; Self et al., 2011) providing high resolution insights into past changes in climate (Brooks and Langdon, 2014), and validation of climate models (Heiri *et al.*, 2014).

1.1 Holocene climate variability

Holocene climate variability is subdued (±2-3^{o}C) (Mayewski *et al*., 2004; O'Brien *et al.*, 1995) compared with the preceding Late Glacial (*c.* 15,000-11,700 years before present [yrs BP], ±7-10^{o}C) (Alley, 2000; Anderson, 1997), nevertheless rapid climate change events are recognised in Holocene palaeoclimate records (Marcott et al., 2013; Mayewski et al., 2004). Changes in insolation caused by solar forcing generally regarded as the dominant driver of global climate change during the Holocene (Mayewski *et al.*, 2004; Wanner *et al*., 2008). The Roman warm period (250 BC-400 AD [2200-1550 yrs BP]), and cooling during the Little Ice Age (LIA) (1350-1850 AD [600-100 yrs BP]) are well established features, notably across the Northern Hemisphere (Johnsen *et al*., 2001; O'Brien *et al*., 1995). Some evidence from the tropics suggests Holocene climate fluctuations such as the LIA maybe global events (Thompson *et al*., 2002; Wanner *et al*., 2008); however, additional quantitative palaeoclimate records are needed to understand the expression of such events in the tropics, and to clarify global climate teleconnections. Although the low latitudes receive 47% of planetary insolation, the climate response in the tropics to solar variability is poorly understood (Crowley, 2000; Polissar *et al.*, 2006).

frazer 2/25/2016 1:41 PM
Deleted: period

frazer 2/25/2016 1:41 PM
Deleted: 25

frazer 2/25/2016 3:44 PM
Deleted: is

frazer 2/25/2016 1:47 PM
Deleted: Growing

frazer 2/25/2016 1:48 PM
Deleted: are probably

1.2 Holocene climate variability in tropical South America

The most notable feature of current South American climate is the annual migration of the Intertropical Convergence Zone (ITCZ), which affects rainfall patterns across the tropical Andes (Garreaud et al., 2009; Hastenrath, 2012). On Holocene timescales, however, there remain large uncertainties regarding the patterns and processes of climate change in the Andes with evidence for both rapid (*c.* 100-1000 yr) precipitation (Haug *et al.*, 2001) and temperature (Thompson *et al.*, 2006; Wanner *et al.*, 2008). A further point to note is the spatial heterogeneity of Holocene climate variability in the tropical Andes (Baker and Fritz, 2015a), particularly regarding precipitation. Ice core records from the Peruvian and Bolivian Andes since *c.* 5400 cal yrs BP suggest the overall trend is towards a drier climate with high amplitude fluctuations and periods of significant aridity. Precipitation reached a minimum during the period between 3800-2800 cal yrs BP and the LIA (Haug *et al.*, 2001; Thompson *et al.*, 1986; Thompson *et al.*, 1995). Speleothem records from the Central Andes of Peru contradict this, however, and indicate instead that from the 15th to 18th century precipitation was on average about 10% higher than the present day (Reuter *et al.*, 2009).

The mid- to late-Holocene (*c.* 6000 cal yrs BP to present) is a period of cooling climate in South America. Pollen evidence suggests montane vegetation replaced Andean forest taxa as the treeline lowered with modern vegetation patterns becoming established by *c.* 3000 cal yrs BP (Markgraf, 1989). Long-term cooling in the late Holocene culminated in a minimum during the 17th and 18th centuries, coinciding with evidence for precipitation minimum during the LIA in northern South America (Haug *et al.*, 2001; Thompson *et al.*, 1986; Thompson *et al.*, 1995). Further south, Patagonian proxy records infer periods that were wet and cold enough to allow glacial advance (Meyer and Wagner, 2008). In the South American tropics, where the relationship between changes in temperature and precipitation are complex (Baker *et al.*, 2001; Garreaud *et al.*, 2009), more independent quantitative estimates of past temperature are needed in order to resolve climate patterns over the tropical Andes during the Holocene.

1.3 Aims

In this study, we have developed the first chironomid-based temperature calibration data set from the tropical Andes (0 to 17°S). Surface sediment samples from 59 lakes along the eastern flank of the Andes to

frazer 2/25/2016 3:44 PM
Deleted: variability

frazer 2/26/2016 2:32 PM
Deleted: ... [1]

Amazonia are analysed. Two approaches are used to develop the inference model, the widely used weighted averaging method (Brooks and Birks, 2000) and a Bayesian approach (Holden *et al*., 2008) which has rarely been used before. The models are applied to fossil chironomid assemblages in a late-Holocene lake sediment record from Laguna Pindo, central Ecuador, to reconstruct mean annual temperature (MAT) changes over the past *c.* 3000 years.

frazer 2/25/2016 3:45 PM
Deleted: a second using

2 Study Sites

2.1 Modern calibration dataset

Surface sediments were collected from 59 lakes across Bolivia (15 lakes), Peru (32 lakes) and Ecuador (12 lakes) between 2004 and 2013 over an altitudinal gradient from 150 m above sea level (a.s.l) to 4655 m a.s.l, between 0-17°S and 64-78°W (Fig 1). The study sites cover an MAT of 25°C; the coldest lake in the data set is 0.8°C MAT and the warmest is 25°C MAT (Table 1). The deepest lake is 25 m and the shallowest is 0.1 m, mean water depth of all the study sites is 5 m. Cold, high elevation lakes are more common within the calibration data set and there are no lakes between 16°C and 20°C. Sediment samples used in this study were taken from the uppermost centimetre (0-1cm) which represents the most recent deposits (approx. 5-20 years) (Frey, 1988) and therefore most comparable with the available climate data for calibration.

frazer 2/25/2016 3:45 PM
Deleted: gradient

frazer 2/26/2016 3:46 PM
Deleted: 1

frazer 2/26/2016 2:37 PM
Deleted: ... [2]

2.2 Fossil chironomid record

Laguna Pindo is a small shallow lake on the eastern flank of the Ecuadorian Andes (1°27.132'S; 78°04.847'W) (Fig1). The site is located at an elevation of 1248 m a.s.l. MAT is *c.* 20°C with little seasonal variation and mean annual precipitation (MAP) can reach *c.* 4000 mm per year (Hijmans *et al*., 2005). Currently the lake is not directly fed by a stream in-flow and has no visible stream out-flow; the lake receives water from surface run-off and direct precipitation. There are no obvious geomorphological causes for the escarpment of the lake and we hypothesise it is tectonic in origin.

At the time of field work (January 2013) maximum water depth was *c.* 1 m, the lake is heavily overgrown with aquatic macrophytes making a detailed bathometric survey difficult. A sedimentary sequence 929 cm long was extracted using a cam-modified piston Livingston corer (Colinvaux *et al.*, 1999) from the centre of the lake to minimise the chance of encountering a sedimentary gap caused by any periods of lake area reductions. Sediments were recovered in aluminium tubes and sealed on site before being transported to the UK and stored at *c.* 4°C. A total of 6 samples were analysed for ^{14}C radiocarbon using AMS dating at the SUERC radiocarbon facility, East Kilbride (Table 2). An age-depth model was created using version 2.2 of the statistical package clam.R (Blaauw, 2010) and the Southern Hemisphere calibration curve SHCal13.14C (Hogg *et al.*, 2013) (Fig 2). The sampling interval for chironomid analysis was not uniform due to a varied sedimentation rate. To achieve as even a coverage possible over the time interval, samples were taken between every 10 and 20cm.

3. Methods

3.1 Chironomid analysis

Chironomid preparation and identification from both lake surface and core sediments followed standard methods as described by Brooks *et al* (2007). The wet sediment was deflocculated in 10% KOH for 2 minutes at 75°C. The sediment was then washed through 212µm and 90µm sieves with water. Chironomids were picked from the residues in a Bogorov counting tray using a stereomicroscope at 25x magnification. Head capsules were mounted in Euparal, ventral side up and identified to the highest possible taxonomic resolution under a compound light microscope at 200-400x magnifications with reference to Wiederholm (1983), Epler (2001) Rieradevall & Brooks (2001), Brooks *et al* (2007), Cranston (2010) and local taxonomic works including Prat *et al* (2011), and Trivinho-Strixino (2011). Some taxa could not be formally identified and so were given informal names. Images and descriptions of informally named taxa are provided in Matthews-Bird *et al* (2015).

3.2 Environmental variables

Environmental variables (depth, pH, conductivity, and water temperature) were measured at each lake in the field. Organic content of the sediment was established through Loss-on-ignition following standard methods as described by Heiri et al (2001). Climate data (MAT, MAP) were obtained from high resolution, interpolated climate surfaces (Hijmans *et al.*, 2005). A summary of all the environmental variables measured can be found in Table 1.

frazer 2/26/2016 3:37 PM
Deleted: 14

frazer 2/26/2016 3:57 PM
Formatted: Tabs: 3.61", Centered

frazer 2/26/2016 2:39 PM
Deleted: Elevation (m a.s.l.), latitude (decimal degrees), and longitude (decimal degrees) were also included as variables in the calibration data set.

3.3 Exploratory statistics

Detrended Correspondence Analysis (DCA) was initially used as an indirect ordination method to assess the gradient lengths in compositional units of taxon turnover (Hill and Gauch, 1980). The gradient length of DCA axis 1 was 5.2 standard deviation units (SD), which suggests a unimodal response, and that linear ordination methods were not appropriate (ter Braak, 1987). Canonical Correspondence Analysis (CCA) was used to explore the influence of the measured environmental variables on the distribution and abundance of taxa. Highly correlated variables were partialled-out by analysis of the variance of their regression coefficients indicated by their Variance Inflator Factors (VIFs). Variables with high VIFs were systematically removed from the environmental variable data set until the remaining variables had VIFs below 20. Detrended canonical correspondence analysis (DCCA) was used to test how much of the variance in the assemblage data was explained by each individual explanatory variable. The ratio of λ_1:λ_2 (i.e., the ratio of first constrained DCCA axis 1 and second unconstrained DCA axis 2) was used to assess the influence an explanatory variable has in describing the variance in the chironomid community assemblage, and hence its predictive power (Juggins, 2013). All taxa were retained in the statistical analysis and rare taxa were down-weighted in the weighted average transfer function (down-weighting of rare species is implicit in the Bayesian approach). Multivariate analysis was carried out on square root transformed chironomid percentage data.

Inference models were developed using two separate approaches. The first method relied on weighted averaging methods, a tried and tested technique well established in quantitative palaeoecology (Birks, 1998; Birks *et al.*, 2012; ter Braak and Juggins, 1993; ter Braak and Looman, 1986). The second method uses a Bayesian approach, which in general has received less attention (Holden *et al.*, 2008). There are a number of inherent problems associated with quantitative inference models (Huntley, 2012; Juggins, 2013; Velle et al., 2010) so the two independent methods were used to compare results and assess the strengths and weaknesses of each method.

The assemblage data was unimodal suggesting transfer functions using weighted averaging partial least squares (WA-PLS) were appropriate (ter Braak and Juggins, 1993). Inference models were also developed using classical and inverse weighted averaging (WA) to compare performance. The optimal number of components

frazer 2/26/2016 2:39 PM
Deleted:

frazer 2/26/2016 2:39 PM
Formatted: Indent: First line: 0"

frazer 2/26/2016 2:41 PM
Formatted: Indent: First line: 0.25"

frazer 2/26/2016 2:41 PM
Deleted:

frazer 2/25/2016 1:24 PM
Deleted: **3.4 Inference models**

frazer 2/25/2016 1:24 PM
Deleted:
... [3]

was assessed using leave-one-out cross validation (jack knifing) and a minimum 5% change in prediction error between components. Sample specific errors for the inferred temperatures were obtained through bootstrapping 999 cycles.

Bayesian model selection was used to generate probability-weighted species response curves (SRCs) for each taxon in the calibration dataset. Each taxon is assigned 8,000 possible SRCs. Each of these SRCs has a probability weight based on its relative ability to describe the training data for that taxon. To perform a reconstruction, likelihood functions (temperature probability distributions) are derived from each taxon in a fossil sample, considering all 8,000 SRCs. Combining the likelihood functions of all the taxa in the fossil sample derives the reconstruction. The power of the Bayesian approach is that it ascribes a probability distribution to the reconstruction, providing a reconstruction-specific uncertainty. An important benefit is that all taxa in the sample provide potentially useful information, even those with low counts that would be largely neglected in a weighted averaging approach. To illustrate, a few counts of a taxon with a narrow temperature tolerance may constrain the Bayesian reconstruction more than a very high count of a taxon with a broad tolerance.

frazer 2/25/2016 1:24 PM
Formatted: Indent: First line: 0"

frazer 2/25/2016 1:24 PM
Deleted: ... [4]

Although the Bayesian model was developed for application to pH reconstructions from diatom assemblages, it is generally applicable whenever it is appropriate to assume a unimodal species response to an environmental gradient. The only modification required is the specification of appropriate priors. The *a priori* probability distribution for optimum temperature in the SRCs was assigned to be uniform in the range -4.2 to +30.8°C (training set range ±°5C). The *a priori* probability for SRC tolerance was assigned to be uniform in the range 2 to 10°C. Other SRC priors were unchanged from those in Holden *et al*. (2008).

DCCA detrending by segments, non-linear rescaling, and constrained by radiocarbon age was used to determine compositional turnover constrained within the stratigraphic sequence (Birks and Birks, 2008). The goodness-of-fit to temperature was evaluated by including the fossil chironomid samples passively in a CCA ordination space of the modern training set samples constrained by MAT. Fossil samples with a squared residual distance within the extreme 10% of the modern calibration dataset samples are considered as having a poor fit to temperature. The modern analogue technique was used to test if fossil samples had good analogues within the modern calibration data set. Any fossil sample with a squared chord distance larger than the 95% threshold of the calibration data set is considered to have no good modern analogues (Birks, 1998; Velle *et al*., 2005).

frazer 2/25/2016 1:24 PM
Deleted: ... [5]

frazer 2/25/2016 2:06 PM
Formatted: Justified

Data were untransformed prior to analysing the dissimilarity using the modern analogue technique. The significance of the final reconstruction was tested by comparing the amount of variance in the fossil data explained by that reconstruction, compared with inferences produced by transfer functions trained on randomly generated environmental data (Telford and Birks, 2011a). In this case, 999 random environmental variables were generated in order to produce the null distribution.

4. Results

4.1 Explanatory variables

The eight remaining explanatory variables, after those with VIFs >20 were removed; together explain 34.03% of the variance (Fig 3). The first two CCA axes explained 61.7% of the variance (λ_1=0.792, λ_2 = 0.466). MAT describes most of the variance in the chironomid assemblages and has the highest λ_1:λ_2 ratio (Table 3). When used as a single explanatory variable, MAT explains 12.93% of the variance (λ_1/λ_2= 1.431).

frazer 2/26/2016 3:57 PM
Deleted: 2

frazer 2/26/2016 3:49 PM
Deleted: 2

4.2 Calibration data set taxa

In total, 55 chironomid taxa were identified in the 59 training set lakes. *Chironomus anthracinus*-type was the most widespread taxon, occurring over the entire temperature gradient (Fig 4). Orthocladiinae are generally most abundant towards the cold end of the temperature gradient. *Cricotopus/Paratrichocladius* type III is the dominant taxon of the coldest lake and is not present in sites >10°C MAT. Figure 4 shows the weighted average and Bayesian optima and tolerance of each taxon ordered by lowest to highest optima as modelled in the weighted averaging approach. In general the temperature optima predicted by each method are similar, however, *Tanytarsus* type II and *Cricotopus/Paratrichocladius* type VII have colder optima when modelled using a Bayesian approach. *Cricotopus/Paratrichocladius* type IV has the coldest temperature optimum *c.* 3.3°C (Fig 5). Few Chironominae were found at the cold end of the calibration data set, but, for example, *Parachironomus* and *Tanytarsus* type II were only found in lakes cooler than *c.* 8°C and had optima of *c.* 7.5°C and *c.* 6.5°C respectively. *Paratanytarsus* and *Pseudosmittia* are important components of the chironomid assemblage between 4-12°C, forming >50% of the chironomid community in some lakes, and have optima of *c.* 9.1 and 8.3°C respectively. *Tanytarsus* type I, *Micropsectra* and *Einfeldia* are dominant taxa at mid-temperatures between *c.*

frazer 2/25/2016 2:22 PM
Deleted: (Matthews-Bird et al., 2015)

frazer 2/26/2016 3:56 PM
Deleted: 3

frazer 2/25/2016 2:22 PM
Deleted: (Matthews-Bird et al., 2015)

frazer 2/25/2016 2:24 PM
Deleted: a

frazer 2/26/2016 3:56 PM
Deleted: 4

10-22°C. The absence of lakes between *c.* 16°C and *c.* 20°C limits a complete understanding of the distribution of taxa occurring at these temperatures.

DCCA analysis, constrained by MAT, indicates an assemblage shift across the temperature gradient of 2.2 SD units. The biggest change in assemblage composition occurs above 12°C MAT (Fig 4). *Goeldichironomus*, *Cladotanytarsus* and *Tanytarsus* type III were only found in lakes with MAT warmer than *c.* 22°C. Tanypodinae were in greatest abundance at the warm end of the temperature gradient between *c.* 10-26°C, *Procladius* was the most common Tanypodinae. It occurred between *c.* 10-26°C and had an optimum of *c.* 21°C.

4.3 Inference models

Both methods (WA and Bayesian) produced similar performance statistics. The best performing model using conventional statistical methods was a WA (inverse) model (Table 4, Fig 6)(R^2_{jack}= 0.890, $RMSEP_{jack}$ = 2.404(°C), Mean $bias_{jack}$= -0.017(°C), Max bias $_{jack}$=4.665(°C)). The Bayesian method produced a slightly higher performing model with R^2_{jack}= 0.909, $RMSEP_{jack}$ =2.373(°C), Mean $bias_{jack}$= 0.598(°C), Max bias $_{jack}$= 3.158(°C).

4.4 Laguna Pindo fossil chironomids and dating

Chironomid remains were found only in the upper 416 cm of the 929cm sequence of Laguna Pindo (Fig 7). In total, 2489 individual chironomid head capsules were analysed. The entire assemblage was made up of 32 taxa in 26 genera and 4 subfamilies. Among the taxa identified, 17 were Chironomini, eight Orthocladiinae and three Tanypodinae. There was high variation between samples both in number of head capsules (average: 82; range: 24 - 184) and concentration per gram of wet sediment (average: 73; range: 2 - 163). There was a marked decline in head capsule concentration below 200 cm. In younger sediments (200-0 cm) head capsule concentration averaged 106/gram, in older samples (200-420 cm) the average was 44/gram. Five zones were identified using optimal partitioning with a broken stick model to define significant zones. *Polypedilum nubifer*-type, *Procladius* and *Limnophyes* were the most abundant taxa; abundances are over 10% wherever they occurred. *Tanytarsus* type II was most abundant below 200 cm (1500 cal yr BP) whilst *Polypedilum nubifer*-type was present in low numbers below 340 cm (2300 cal yr BP). During periods of low *Polypedilum nubifer*-type abundance, *Tanytarsus* type II and *Tanytarsus* type I occur in greater numbers (e.g. 420-360; 290-250 cm).

frazer 2/26/2016 3:56 PM
Deleted: 3

frazer 2/25/2016 2:29 PM
Moved down [1]: Chironomid larval head capsule concentrations can vary significantly between lakes, due to differences in preservation or abundance. Low counts can have adverse effects on the performance of inference models and the reliability of quantitative environmental reconstructions when using conventional methods (Heiri and Lotter, 2001; Quinlan and Smol, 2001). A minimum count size of 50 head capsules per sample is advised (Heiri and Lotter, 2001; Quinlan and Smol, 2001), however, good model performance has been achieved even when several samples include as few as 15-30 head capsules (Massaferro *et al.*, 2014). In some lakes in the current training set head capsule concentrations were as low as two head capsules per gram of sediment. Fifteen lakes in the data set produced fewer than 50 head capsules, and three lakes had fewer than 30. On average 77 individuals were analysed from each lake with a minimum count of 23 and a maximum of 164 (Table 1). Lakes with low head capsule counts were retained in the model in order to maintain as even coverage as possible across the temperature gradient.

frazer 2/26/2016 3:50 PM
Deleted: 3

frazer 2/26/2016 3:56 PM
Deleted: 5

frazer 2/26/2016 2:47 PM
Deleted:

frazer 2/26/2016 3:55 PM
Deleted: 6

The best-fit age depth model for Laguna Pindo was a smooth spline (Fig 2). Due to the absence of chironomids at the bottom of the sequence, six radiocarbon samples were used for building the model with a total depth of the sediment considered of 461 cm (Table 2). The sedimentation rate ranged between 0.03 and 0.5 cm/yr, with a sampling interval resolution of 97 years between samples on average (range from 27 to 196 years).

frazer 2/26/2016 4:00 PM
Deleted: S1

frazer 2/26/2016 4:00 PM
Deleted: S1

4.5 Palaeotemperature reconstruction

Both transfer functions (WA inverse and Bayesian) show similar patterns in the temperature reconstruction (Fig 8). From 3000-2500 cal yr BP inferred temperatures are cold relative to the modern (20.2^{o}C). The minimum WA inverse temperatures are much colder (13.5°C±2.5) than the inferred Bayesian temperatures (17.5°C±3.7) for the early section of the sequence. From 2400 to 1700 cal yr BP inferred temperatures from both methods oscillate around *c.* 18-19°C but remained depressed relative to the modern. A notable feature of both reconstructions is the sudden drop in inferred temperatures at 1600 cal yr BP. Inferred temperatures fall by *c.* 2°C to 17.5°C±2.7. This abrupt drop in temperature is short-lived in both reconstructions and temperatures return to previous values in the subsequent sample. From 1500 cal yr BP to the present the chironomid-inferred temperatures stabilise and steadily rise. Peak temperatures for the entire record (21.9°C±3.5) are inferred between 400-700 cal yr BP. Temperatures begin to cool from 400 cal yrs BP in both reconstructions, reaching a minimum of *c.* 17°C±2.5 *c.* 100 cal yr BP before rising rapidly to between 20-21°C±2.5 in the most recent sediment sample. On average the Bayesian model infers warmer temperatures than the WA model.

frazer 2/26/2016 3:55 PM
Deleted: 7

The fossil samples of Laguna Pindo plot within the modern variation of chironomid assemblages when included passively in a CCA analysis of the calibration data set (Fig 9). This suggests that the calibration dataset is appropriate for the fossil sequence of Laguna Pindo. The fossil samples plot along the MAP gradient suggesting precipitation is an important variable controlling the variance in the fossil assemblages. The sites associated with high precipitation in the calibration dataset are located in the same region of the Ecuadorian Andes as the fossil site. With a modern MAT of *c.* 20°C, however, Laguna Pindo is located in a region of the temperature gradient that is poorly covered in the calibration dataset (Fig 4). Seven taxa found in the Laguna Pindo sequence do not occur in any of the analysed calibration data set lakes. These include three unknown

frazer 2/26/2016 3:55 PM
Deleted: 8

frazer 2/26/2016 3:56 PM
Deleted: 3

frazer 2/25/2016 1:06 PM
Deleted: ,

frazer 2/25/2016 1:06 PM
Deleted: although the samples plot within the range of modern calibration lakes that lie at similar elevations (1000-3000 m a.s.l).

morphotypes, three *Xestochironomus* morphotypes, and *Metriocnemus eurynotus*-type. These taxa, however, never comprise more than 10% of the chironomid assemblage of any one sample.

Fourteen of the fossil samples are considered to have a poor goodness-of-fit to temperature and all fossil samples are considered as having poor modern analogues in the calibration data set (Fig 10). Although the modern analogue technique is not used to infer past temperatures the lack of modern analogues in the fossil assemblage is important when considering the reliability of any reconstruction.

frazer 2/25/2016 2:56 PM
Deleted: Seven

frazer 2/26/2016 3:54 PM
Deleted: 9

DCCA constrained by radiocarbon age shows an abrupt change at 1475 cal yr BP between zones 3 and 4 and a turnover of 1.6 SD units over the whole sequence (Fig 10). Much of the variation in goodness-of-fit and DCCA sample scores is mirrored by changes in count size and head capsule concentration. The sudden drop in head capsule concentration occurs at a step change in DCCA assemblage variation (1475 cal yrs BP) (Fig 10). Periods of increased count size and head capsule concentration in older sediments (2100-2250 cal yrs BP) also coincides with periods of improved goodness-of-fit (Fig 10). The WA classical inferred MAT values using the modern calibration data set explain more of the variance than 95% of randomly generated variables and so the WA classical MAT reconstructions can be deemed statistically significant (*p*= 0.032) (Fig 11) (Telford and Birks, 2011a).

frazer 2/26/2016 3:55 PM
Deleted: 9

frazer 2/26/2016 2:50 PM
Deleted: The most recent sample is clearly distinct from any other period of the record.

frazer 2/26/2016 3:55 PM
Deleted: 9

frazer 2/26/2016 3:55 PM
Deleted: 9

frazer 2/26/2016 3:54 PM
Deleted: 0

5. Discussion

5.1 Chironomids and environmental variables

Chironomids have been shown to respond to temperature at a variety of spatial scales and taxonomic levels (Brooks, 2006; Eggermont and Heiri, 2011). Temperature is a key variable in controlling chironomid development at all stages of their life cycles, and influences voltinism, behaviour and metabolism (Armitage *et al*., 1995). Across the Northern Hemisphere, over large temperature gradients, mean July air temperature, the warmest month of the year, which reflects the developmental period of most species, has been shown to be the major determinant of variation in chironomid assemblages (Brooks, 2006; Walker and Cwynar, 2006). As a result, many quantitative temperature inference models have been developed to reconstruct mean July air temperature. Across the tropics however seasonal variation is small and many chironomids are multivoltine (Walker and Mathews, 1987) so temperatures throughout the year are likely to be relatively more influential. In tropical East Africa, Eggermont *et al*. (2010) demonstrated that mean annual air temperature was a significant driver of chironomid assemblage composition and developed a chironomid-based inference model on this basis. Similarly, Wu *et al*. (2014) showed MAT to be the most important environmental variable when developing a chironomid inference model for Central America. When attempting to make quantitative inferences from fossil assemblages it is first crucial to establish that the variable of interest is an important ecological determinant. The variable to be reconstructed must describe a statistically important component of the variance within the assemblage data (Juggins, 2013). Compared to other measured variables, mean annual temperature explained the largest amount of chironomid assemblage variance and had the highest eigenvalue ratio (λ_1:λ_2) in the Andean calibration dataset (Table 3). The explanatory strength of temperature in the calibration data set meets the minimum criterion proposed by Juggins (2013) (i.e. λ_1:$\lambda_2 > 1.0$) for temperature being a suitable variable to reconstruct from this calibration dataset.

frazer 2/26/2016 3:49 PM
Deleted: 2

The DCCA results suggest that precipitation is also a strong ecological determinant (λ_1:λ_2=0.9); the passive plot of fossil samples with calibration samples further supports this conclusion. The fossil samples of Laguna Pindo are strongly associated with MAP. Precipitation in Andean landscapes, however, is spatially heterogeneous and geographically close localities experience significantly different rainfall patterns (Garreaud *et al*., 2009). Lakes associated with high rainfall (Fig 3) are actually in areas of the northern Andes with two rainy seasons a year. It is very likely that the bimodality of rainfall in these areas is as important in controlling chironomid populations as the total amount of rainfall as measured by MAP. Precipitation is also intrinsically linked to temperature as both temperature and precipitation increase with decreasing latitude in tropical South America (Garreaud *et al*., 2009). Unlike temperature, precipitation affects chironomids indirectly making any quantitative inference difficult. Precipitation will alter a suite of environmental variables (e.g. pH, conductivity, depth, substrate) making quantitative inferences of precipitation problematic. As chironomid life cycles are strongly controlled by temperature and many tropical chironomid species tend to be multivoltine, we suggest the most appropriate variable both ecologically and statistically to reconstruct using the Andean calibration data sets is MAT although the of influence of precipitation cannot be overlooked.

frazer 2/26/2016 3:57 PM
Deleted: 2

The optima and temperature tolerances (Fig 5) of many taxa found in the current study are similar to that noted in other Neotropical chironomid calibration datasets, further supporting the conclusion of

frazer 2/26/2016 3:56 PM
Deleted: 4

temperature being an important ecological determinant. For example, Wu *et al*. (2014) in Central America, found taxa of the genera *Beardius*, *Labrundinia* and *Goeldichironomus* to have optima between 23-24°C whilst *Limnophyes* and *Corynoneura* where more abundant at the colder end of the gradient with optima of 15°C and 18°C respectively. In the current dataset *Beardius, Labrundinia*, and *Goeldichironomus* all have optima between 23-24°C and *Limnophyes* and taxa of *Corynoneura* also have optima of 15°C and 19°C, respectively. *Limnophyes* also has one of the broadest tolerances of all taxa in both calibration datasets suggesting the genus is probably represented by many species (Matthews-Bird et al., 2015). More work is needed in order to refine chironomid larval taxonomy in South America, however the current data suggest the potential for a larger calibration dataset applicable to wider area incorporating the Northern Neotropics and Central America.

5.2 Model performance

Although both models (WA inverse and Bayesian) perform well (WA RMSEP= 2.4°C/ 9.6% of training set range and Bayesian RMSEP= 2.3°C/9.2% of training set range), some of the best performing chironomid-based temperature inference models have prediction errors closer to 1.0°C (Brooks and Birks, 2001; Heiri *et al*., 2011, 2007; Olander *et al*., 1999). The highest performing chironomid inference models often have in excess of 100-150 calibration sites compared with just 59 in the current model and this may account for its reduced performance. Furthermore the lakes in the calibration data set are not evenly distributed over the temperature gradient. The cold end of the gradient has a higher number of lakes (34 cold, high elevation lakes) than at warm and intermediate temperatures (15 warm, mid-low elevation lakes). Uneven sampling has been shown to lead to biases which may reduce RMSEP (Telford and Birks, 2011b). Furthermore the over-representation of cold lakes in the current dataset may result in under-estimation of the temperature optima of some taxa and, therefore, bias temperature estimates towards cold values. In the Andean dataset, as analysis of residuals shows, temperatures around 10°C are often under-estimated (Fig 6). Furthermore, the inferred temperatures of Laguna Pindo are on average cooler than the modern day conditions.

frazer 2/26/2016 3:56 PM
Deleted: 5

The absence of lakes in part of the temperature gradient may limit the reliability of estimates of optima and tolerances of taxa and also create 'edge effects' in the middle of the temperature range, in addition to those that occur at the cold and warm end of the temperature gradient (Eggermont *et al*., 2010). Such problems are inherent to WA models as predicted values are pulled towards the mean of the training set resulting in under- and over-estimations of high and low values (ter Braak and Juggins 1993). However, despite having no lakes between 16-20°C in the calibration data set, additional edge effects are not a feature of the current inference model. The gap of c. 4°C does not appear to have compromised model performance, probably as the interval is not significant and taxa have tolerances that span these temperatures.

Chironomid larval head capsule concentrations can vary significantly between lakes, due to differences in preservation or abundance. Low counts can have adverse effects on the performance of inference models and the reliability of quantitative environmental reconstructions when using conventional methods (Heiri and Lotter, 2001; Quinlan and Smol, 2001). A minimum count size of 50 head capsules per sample is advised (Heiri and Lotter, 2001; Quinlan and Smol, 2001), however, good model performance has been achieved even when several samples include as few as 15-30 head capsules (Massaferro *et al*., 2014). In some lakes in the current training set head capsule concentrations were as low as two head capsules per gram of sediment. Fifteen lakes in the data set produced fewer than 50 head capsules, and three lakes had fewer than 30. On average 77 individuals were analysed from each lake with a minimum count of 23 and a maximum of 164 (Table 1). Lakes with low head capsule counts were retained in the model in order to maintain as even coverage as possible across the temperature gradient.

Polypedilum nubifer-type and *Chironomus anthracinus*-type make up a large component of the chironomid assemblages in lakes across the entire temperature gradient (Fig 4). Such eurythermic taxa probably include several different species. It is difficult to model reliable, or even meaningful, optima for eurythermic taxa. Poor model performance or unreliable reconstructions may result if the assemblage is dominated by eurythermic taxa. We note that eurythermic taxa are described by high tolerance SRCs in the Bayesian approach, leading to increased uncertainty in reconstructions through broad likelihood functions that contribute little information to the posterior. Inferred temperature of *c.* 10°C, are likely to be underestimated as many taxa found at these temperatures also occur in cold lakes, which are over-represented in the calibration data-set. In African lakes Eggermont *et al*. (2010) found that the presence of eurythermic taxa such as *Chironomus* type Kibos caused an overestimation of temperatures in lakes at the warm end of the gradient.

frazer 2/25/2016 2:29 PM
Moved (insertion) [1]

frazer 2/26/2016 3:56 PM
Deleted: 3

They also found that the occurrence of *Limnophyes minimus*-type and *Paraphaenocladius* type OI Bolossat overestimated the temperature of lakes close to where gaps occurred in the gradient (Eggermont *et al*., 2010). Similarly, in a New Zealand calibration data set developed by Woodward and Shulmeister (2006), *Chironomus* was present in both high elevation, cold, oligotrophic lakes and lower elevation, warm, eutrophic lakes. The intermediate temperature optimum estimated for this taxon resulted in over-estimated temperatures of cold lakes and under-estimates of warm lakes (Woodward and Shulmeister, 2006). Eurythermic taxa may be contributing to the over-estimation of cold temperatures and the under-estimation of temperatures in the middle of the gradient in the Andean inference model.

5.3 WA vs Bayesian

Despite similar performance statistics between the Bayesian and WA methods, the inferred pattern of late-Holocene temperature change is different. Temperatures inferred c. 2700 cal yr BP (400 cm) (Fig 8) using the WA inverse method is extremely cold (*c.* 14°C) compared with the rest of the record. This reconstruction is driven by the high abundance of *Tanytarsus* type II, a taxon that has a WA temperature optimum of 6.5°C. The Bayesian reconstruction for this sample of 17.8 ±2.8°C, is in line with more modest temperature shifts that would be expected in the late-Holocene (Wanner *et al*., 2008). One advantage of the Bayesian methodology is the transparency of the reconstruction through consideration of individual likelihood functions for this assemblage (Fig 12). Although *Tanytarsus* type II is abundant in the sample its influence in the reconstruction is moderated by several other taxa with higher temperature optima that are present at low abundances. This temperature estimate demonstrates the Bayesian reconstruction can be sensitive to a few counts of a species that have a negligible effect in a WA approach. The likelihood function for Chironomini type II, which has an abundance of only 2.3% in the sample, constrains the reconstruction more than *Tanytarsus* type II, which has an abundance of 74%. This is because Chironomini type II is only found in the warmest lakes in the calibration set, each time with a low abundance. We note that because it is found in only three training set sites, Chironomini type II is associated with many (671) high-probability SRCs, defined as having a probability great that 10% of the most likely SRC. For this reason, its likelihood function is relatively broad and extends to temperatures far lower than the temperature of the sites in which the taxon is found in the training set.

frazer 2/26/2016 3:55 PM
Deleted: 7

frazer 2/26/2016 3:54 PM
Deleted: 1

5.4 Temperature and secondary environmental variables

Whilst the λ1/λ2 of 1.431 indicates that MAT is appropriate for reconstruction using this calibration dataset (Juggins, 2013), it does not necessarily mean that reliable temperature reconstructions can be obtained from a fossil record (Telford and Birks, 2011a). Before attempting to interpret any reconstruction several metrics can be used to assess the validity of a reconstruction (Juggins and Telford, 2012).

The modern analogue technique compares the similarity of the fossil samples to the modern samples in the calibration data set. All fossil samples are greater than the 5th percentile of the square chord distance (Fig 10), which suggests there is no close modern analogue in the calibration set to any fossil sample (Birks, 1998; Juggins and Birks, 2001). The lack of modern analogues in the Laguna Pindo fossil sequence is due to the many taxa present in the fossil samples that are not present in the calibration data set. This may reflect the lack of lakes in the calibration dataset with MAT values close to those of Laguna Pindo. Nevertheless, WA and WAPLS models have been shown to perform well in non-analogue situations (Birks *et al.*, 2010). The Bayesian method generates temperature reconstructions from likelihood functions of species in the calibration data set. Although analogous assemblages are not required for the Bayesian reconstruction (each taxon is treated equally and individually), species that are absent from the training set cannot contribute information to the posterior, thereby increasing the uncertainty associated with the reconstruction. One advantage of the Bayesian methodology is that this uncertainty is explicitly incorporated into the Bayesian reconstruction (Holden *et al.*, 2008).

During periods of poor fit-to-temperature, variables other than temperature may have been affecting the composition of the chironomid assemblage. As noted previously, the CCA biplot of fossil samples included passively with the significant explanatory variables (Fig 9) shows that MAP was also important in driving the assemblage variance. During times of poor fit to temperature the influence of precipitation as a secondary variable may be more important than temperature in influencing the chironomid assemblage composition. Indeed, precipitation has been shown to be an important variable in controlling the modern distribution of chironomid taxa in the tropical Andes (Matthews-Bird et al., 2015).

frazer 2/26/2016 3:01 PM
Deleted: ... [6]

frazer 2/26/2016 3:01 PM
Deleted: 5

frazer 2/26/2016 3:54 PM
Deleted: 9

frazer 2/26/2016 3:55 PM
Deleted: 8

Samples with poor fit-to-temperature also corresponded with samples having low numbers of head capsules. The number of head capsules retrieved will directly affect how representative a sample is to the chironomid fauna (Heiri, 2004; Quinlan and Smol, 2001). The cold oscillations inferred from the Bayesian reconstruction are more in line with what is expected during the late-Holocene (1-3°C); the likelihood functions of rare species, which favour warm conditions, combine to rule out the anomalously cold temperatures suggested by some of the WA reconstructions. As discussed above, the over-representation of cold lakes in the calibration dataset will likely bias species optima to colder values in a weighted average approach so there may be a tendency for the model to underestimate temperature, especially during cold periods. This problem is likely exaggerated when head capsule concentration is low, cold indicator taxa may have higher abundances than would be the case if all taxa were accurately represented.

The DCCA results indicate that there was a distinct change in the composition of the chironomid assemblage after 1600 cal yr BP (210 cm). This largely coincides with an increase in head capsule concentration, possibly indicating an increase in lake productivity, and the shift in chironomid-inferred temperatures from low to high. Indeed post 1600 cal yr BP, (210 cm) samples are inferred as being on average 2-3°C warmer than early sections using Bayesian and WA models respectively.

Although the temperature reconstruction has a good ecological basis, because chironomids globally are highly sensitive to temperature and, Laguna Pindo is on an ecotonal boundary that is sensitive to temperature changes, precipitation is influential as a secondary variable. The WA inverse MAT reconstruction, however, is statistically significant based on the criteria described by Telford and Birks (2011a) (Fig 11) suggesting that despite precipitation as a possible confounding variable, a temperature signal can be obtained from Neotropical chironomids. We would caution, however, against an over interpretation at this stage. Due to some of the limitations discussed previously, the reconstruction can only currently be deemed qualitative and requires more research before quantitative estimates can be generated with confidence.

frazer 2/26/2016 3:54 PM
Deleted: 0

frazer 2/26/2016 3:11 PM
Formatted: Indent: First line: 0"

6. Conclusions

frazer 2/26/2016 3:11 PM
Deleted: 5.6 Cooling climate 3800-2800 ca ... [7]

The chironomid fauna of the tropical Andes have been shown to be sensitive to climate variables, particularly temperature and precipitation. Both variables (MAT and MAP) meet the basic criteria for being used in an environmental reconstruction using the Andean calibration dataset. MAT, however, is an important determinant of chironomid species distribution and abundance and was therefore more appropriate to be reconstructed. The influence of precipitation should be explored further and must be considered as an important secondary variable especially when reconstructing past conditions in the region. It is very likely that the influence of precipitation noted here relates to the annual variability in rainfall across the Andes as opposed to overall amount making any quantitative interpretations even more difficult.

The two techniques used to develop inference models (WA and Bayesian) show comparable performance statistics (WA inverse model R^2_{jack}= 0.890, $RMSEP_{jack}$ = 2.404(°C), Mean $bias_{jack}$= -0.017(°C), Max bias $_{jack}$=4.665(°C); Bayesian model R^2_{jack}= 0.909, $RMSEP_{jack}$ =2.373(°C), Mean $bias_{jack}$= 0.598(°C), Max bias $_{jack}$= 3.158(°C)). This work demonstrates a proof of method, however, a larger calibration dataset with a more even coverage of calibration sites is needed in order to improve model performance. The Bayesian approach provided a transparent reconstruction less susceptible to the effect of an uneven distribution of calibration sites and performed particularly well during periods of low count size and when inferring cold intervals. The chironomid-based MAT reconstruction from the Laguna Pindo is often colder than would be expected for Holocene timescales. The underestimated temperatures are most likely the direct result of an over representation of cold lakes in the calibration dataset. The addition of more calibration sites between 12°C and 20°C would expand our understanding of tropical Andean chironomid distribution, and significantly improve model performance and reconstruction reliability.

Knowledge of past tropical climate dynamics is fundamental not only to understanding regional climate but also global climate patterns and hemispherical teleconnections. Quantitative temperature proxies, such as chironomids, will provide valuable data on past climate variability in the region. The reconstructions presented here demonstrate the potential of the proxy and also highlights the complexity of late-Holocene climate change in tropical South America.

frazer 2/26/2016 3:07 PM
Deleted: fossil record suggests that periods of low solar output not only affect the tropics through changes in precipitation, but also directly affect tropical temperatures. Inferred temperatures were 2-3°C cooler relative to the modern during the widely recognised 3500-2500 cal yr BP cooling event. Long-term cooling during the late Holocene is not apparent in the Laguna Pindo record. However, temperatures do cool by 1-2.2°C relative to the modern during the LIA period, although this is only noted in a single fossil sample.

Acknowledgements

Funding was provided by the Natural Environment Research Council (NERC), UK. NERC grant (ref: NE/J018562/1) was awarded to E. Montoya and (ref: NE/J500288/1) awarded to F. Matthews-Bird. This work was supported by the NERC Radiocarbon Facility NRCF010001 (allocation number 1682.1112). Special thanks to Dr Pauline Gulliver for her continuous involvement and support during radiocarbon dating. The authors also wish to thank Mark Bush, Francis Mayle, Yarrow Axford, Alex Chepstow-Lusty and Mick Frogley for their kind donation of samples.

References

Alley, R.B., 2000. The Younger Dryas cold interval as viewed from central Greenland. Quat. Sci. Rev. 19, 213–226.

Anderson, E.D., 1997. Younger Dryas research and its implications for understanding abrupt climatic change. Prog. Phys. Geogr. 21, 230–249.

Armitage, P.D., Cranston, P.S., Pinder, L.C.V., 1995. The Chironomidae: the biology and ecology of nonbiting midges. London: Chapman and Hall.

Baker, P. a., Fritz, S.C., 2015b. Nature and causes of Quaternary climate variation of tropical South America. Quat. Sci. Rev. 124, 31–47. doi:10.1016/j.quascirev.2015.06.011

Baker, P.A., Seltzer, G.O., Fritz, S.C., Dunbar, R.B., Grove, M.J., Tapia, P.M., Cross, S.L., Rowe, H.D., Broda, J.P., 2001. The history of South American tropical precipitation for the past 25,000 years. Science 291, 640–643. doi:10.1126/science.291.5504.640

Bird, B.W., Abbott, M.B., Vuille, M., Rodbell, D.T., Stansell, N.D., Rosenmeier, M.F., 2011. A 2,300-year-long annually resolved record of the South American summer monsoon from the Peruvian Andes. Proc. Natl. Acad. Sci. U. S. A. 108, 8583–8. doi:10.1073/pnas.1003719108

Birks, H.J.B., 1998. Numerical tools in palaeolimnology–Progress, potentialities and problems. J. Paleolimnol. 20, 307–332.

Birks, H.J.B., Birks, H.H., 2008. Biological responses to rapid climate change at the Younger Dryas Holocene transition at Krakenes, western Norway. The Holocene 18, 19–30. doi:10.1177/0959683607085572

Birks, H.J.B., Heiri, O., Seppä, H., Bjune, A.E., 2010. Strengths and Weaknesses of Quantitative Climate Reconstructions Based on Late-Quaternary Biological Proxies. Open Ecol. J. 3, 68–110.

Birks, H.J.B., Lotter, A.F., Juggins, S., Smol, J.P. (Eds.), 2012. Tracking Environmental Change Using Lake Sediments; Data Handling and Numerical Techniques. Springer Netherlands.

Blaauw, M., 2010. Methods and code for “classical” age-modelling of radiocarbon sequences. Quat. Geochronol. 5, 512–518. doi:10.1016/j.quageo.2010.01.002

frazer 2/25/2016 3:28 PM

Deleted: Baker, P. a., Fritz, S.C., 2015a. Nature and causes of Quaternary climate variation of tropical South America. Quat. Sci. Rev. 124, 31–47. doi:10.1016/j.quascirev.2015.06.011

frazer 2/25/2016 3:29 PM

Deleted: -

Brooks, S.J., 2006. Fossil midges (Diptera: Chironomidae) as palaeoclimatic indicators for the Eurasian region. Quat. Sci. Rev. 25, 1894–1910. doi:10.1016/j.quascirev.2005.03.021

Brooks, S.J., 2000. Chironomid-inferred Late-glacial air temperatures at Whitrig Bog, Southeast Scotland. J. Quat. Sci. 15, 759–764.

Brooks, S.J., Axford, Y., Heiri, O., Langdon, P.G., Larocque-Tobler, I., 2012. Chironomids can be reliable proxies for Holocene temperatures. A comment on Velle et al. (2010). The Holocene 22, 1495–1500. doi:10.1177/0959683612449757

Brooks, S.J., Birks, H.J.B., 2001. Chironomid-inferred air temperatures from Lateglacial and Holocene sites in north-west Europe: progress and problems. Quat. Sci. Rev. 20, 1723–1741. doi:10.1016/S0277-3791(01)00038-5

Brooks, S.J., Birks, H.J.B., 2000. Chironomid-inferred late-glacial and early-Holocene mean July air temperatures for Kråkenes Lake, western Norway. J. Paleolimnol. 23, 77–89.

Brooks, S.J., Langdon, P.G., 2014. Summer temperature gradients in northwest Europe during the Lateglacial to early Holocene transition (15–8 ka BP) inferred from chironomid assemblages. Quat. Int. 1–11. doi:10.1016/j.quaint.2014.01.034

Brooks, S.J., Langdon, P.G., Heiri, O., 2007. The Identification and use of Palaearctic Chironomidae Larvae in Palaeoecology. QRA Technical Guide No. 10, Quaternary Research Association, London.

Čiamporová-Zaťovičová, Z., Hamerlík, L., Šporka, F., Bitušík, P., 2010. Littoral benthic macroinvertebrates of alpine lakes (Tatra Mts) along an altitudinal gradient: a basis for climate change assessment. Hydrobiologia 648, 19–34. doi:10.1007/s10750-010-0139-5

Colinvaux, P., De Oliveira, P.E., Patino, J.E.., 1999. Amazon Pollen Manual and Atlas. Harwood Academic Publishers.

Collins, M., An, S.-I., Cai, W., Ganachaud, A., Guilyardi, E., Jin, F.-F., Jochum, M., Lengaigne, M., Power, S., Timmermann, A., Vecchi, G., Wittenberg, A., 2010. The impact of global warming on the tropical Pacific Ocean and El Niño. Nat. Geosci. 3, 391–397. doi:10.1038/ngeo868

Cranston, P.S., 2010. URL http://chirokey.skullisland.info/.

Crowley, T.J., 2000. Causes of Climate Change Over the Past 1000 Years. Science, 289, 270–277. doi:10.1126/science.289.5477.270

Dimitriadis, S., Cranston, P.S., 2001. An Australian Holocene climate reconstruction using Chironomidae from a tropical volcanic maar lake. Palaeogeogr. Palaeoclimatol. Palaeoecol. 176, 109–131.

Eggermont, H., Heiri, O., 2011. The chironomid-temperature relationship: expression in nature and palaeoenvironmental implications. Biol. Rev. Camb. Philos. Soc. 87, 430–456. doi:10.1111/j.1469-185X.2011.00206.x

Eggermont, H., Heiri, O., Russell, J., Vuille, M., Audenaert, L., Verschuren, D., 2010. Paleotemperature reconstruction in tropical Africa using fossil Chironomidae (Insecta: Diptera). J. Paleolimnol. 43, 413–435. doi:10.1007/s10933-009-9339-2

Epler, J.H., 2001. Identification manual for the Larval Chironomidae (Diptera) of South Carolina.

Frey, D.G., 1988. Littoral and offshore communities of diatoms, cladocerans and dipterous larvae, and their interpretation in paleolimnology. J. Paleolimnol. 1, 179–191.

Garreaud, R.D., Vuille, M., Compagnucci, R., Marengo, J., 2009. Present-day South American climate. Palaeogeogr. Palaeoclimatol. Palaeoecol. 281, 180–195. doi:10.1016/j.palaeo.2007.10.032

Hastenrath, S., 2012. Climate dynamics of the tropics. Springer Science & Business Media.

Haug, G.H., Hughen, K.A., Sigman, D.M., Peterson, L.C., Röhl, U., 2001. Southward migration of the intertropical convergence zone through the Holocene. Science 293, 1304–8. doi:10.1126/science.1059725

Heiri, O., 2004. Within-lake variability of subfossil chironomid assemblages in shallow Norwegian lakes. J. Paleolimnol. 32, 67–84. doi:10.1023/B:JOPL.0000025289.30038.e9

frazer 2/25/2016 3:30 PM
Deleted: No Title [WWW Document].

frazer 2/25/2016 3:30 PM
Deleted: (80-.)

Heiri, O., Brooks, S.J., Birks, H.J.B., Lotter, A.F., 2011. A 274-lake calibration data-set and inference model for chironomid-based summer air temperature reconstruction in Europe. Quat. Sci. Rev. 30, 3445–3456. doi:10.1016/j.quascirev.2011.09.006

Heiri, O., Brooks, S.J., Renssen, H., Bedford, A., Hazekamp, M., Ilyashuk, B., Jeffers, E.S., Lang, B., Kirilova, E., Kuiper, S., Millet, L., Samartin, S., Toth, M., Verbruggen, F., Watson, J.E., van Asch, N., Lammertsma, E., Amon, L., Birks, H.H., Birks, H.J.B., Mortensen, M.F., Hoek, W.Z., Magyari, E., Muñoz Sobrino, C., Seppä, H., Tinner, W., Tonkov, S., Veski, S., Lotter, A.F., 2014. Validation of climate model-inferred regional temperature change for late-glacial Europe. Nat. Commun. 5, 4914. doi:10.1038/ncomms5914

Heiri, O., Cremer, H., Engels, S., Hoek, W.Z., Peeters, W., Lotter, A.F., 2007. Lateglacial summer temperatures in the Northwest European lowlands: a chironomid record from Hijkermeer, the Netherlands. Quat. Sci. Rev. 26, 2420–2437. doi:10.1016/j.quascirev.2007.06.017

Heiri, O., Lotter, A., 2001. Effect of low count sums on quantitative environmental reconstructions: an example using subfossil chironomids. J. Paleolimnol. 26, 343–350.

Heiri, O., Lotter, A.F., Hausmann, S., Kienast, F., 2003. A chironomid-based Holocene summer air temperature reconstruction from the Swiss Alps. The Holocene 13, 477–484.

Heiri, O., Lotter, A.F., Lemke, G., 2001. Loss on ignition as a method for estimating organic and carbonate content in sediments: reproducibility and comparability of results. J. Paleolimnol. 25, 101–110.

Hijmans, R.J., Cameron, S.E., Parra, J.L., Jones, P.G., Jarvis, A., 2005. Very high resolution interpolated climate surfaces for global land areas. Int. J. Climatol. 25, 1965–1978. doi:10.1002/joc.1276

Hill, M., Gauch, H.., 1980. Detrended correspondence analysis: an improved ordination technique. Vegetatio 42, 47–58.

Hogg, A.G., Hua, Q., Blackwell, P.G., Niu, M., Buck, C.E., Guilderson, T.P., Heaton, T.J., Palmer, J.G., Reimer, P.J., Reimer, R.W., Turney, C.S.M., Zimmerman, S.R.H., 2013. SHCal13 Southern Hemisphere calibration, 0–50,000 cal yr BP. Radiocarbon 55, 1889–1903.

Holden, P.B., Mackay, A.W., Simpson, G.L., 2008. A Bayesian palaeoenvironmental transfer function model for acidified lakes. J. Paleolimnol. 39, 551–566. doi:10.1007/s10933-007-9129-7

Huntley, B., 2012. Reconstructing palaeoclimates from biological proxies: Some often overlooked sources of uncertainty. Quat. Sci. Rev. 31, 1–16. doi:10.1016/j.quascirev.2011.11.006

Ivanochko, T., Ganeshram, R., Brummer, G., Ganssen, G., Jung, S., Moreton, S., Kroon, D., 2005. Variations in tropical convection as an amplifier of global climate change at the millennial scale. Earth Planet. Sci. Lett. 235, 302–314. doi:10.1016/j.epsl.2005.04.002

Jansen, E., Overpeck, J., Briffa, K.., Duplessy, J.-C., Joos, F., Masson-Delmotte, V., Olago, D., Otto-Bliesner, B., Peltier, W.., Rahmstorf, S., Ramesh, R., Raynaud, D., Rind, O., Solomina, O., Villalba, R., Zhang, D., 2007. Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovvernmental Panel on Climate Change, in: Solomon, S., Qin, D., Manning, M., Chen, Z., Marquis, M., Averyt, K.B., Tignor, M., Miller, H.. (Eds.), . Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Johnsen, S.J., Dahl-Jensen, D., Gundestrup, N., Steffensen, J.P., Clausen, H.B., Miller, H., Masson-Delmotte, V., Sveinbjornsdottir, A.E., White, J., 2001. Oxygen isotope and palaeotemperature records from six Greenland ice-core stations: Camp Century, Dye-3, GRIP, GISP2, Renland and NorthGRIP. J. Quat. Sci. 16, 299–307. doi:10.1002/jqs.622

Jomelli, V., Favier, V., Rabatel, A., Brunstein, D., Hoffmann, G., Francou, B., 2009. Fluctuations of glaciers in the tropical Andes over the last millennium and palaeoclimatic implications: A review. Palaeogeogr. Palaeoclimatol. Palaeoecol. 281, 269–282. doi:10.1016/j.palaeo.2008.10.033

Jones, P., Mann, M., 2004. Climate over past millennia. Rev. Geophys. 42, 1–42. doi:10.1029/2003RG000143.CONTENTS

Juggins, S., 2013. Quantitative reconstructions in palaeolimnology: new paradigm or sick science? Quat. Sci. Rev. 64, 20–32. doi:10.1016/j.quascirev.2012.12.014

frazer 2/25/2016 3:32 PM
Deleted: e

Juggins, S., Birks, H.J.B., 2001. Quantitative Environmental Reconstructions from Biological Data, in: Birks, H.J.B., Lotter, A.F., Juggins, S., Smol, J.P. (Eds.), Tracking Environmental Change Using Lake Sediments, Developments in Paleoenvironmental Research 5. pp. 431–494.

Juggins, S., Telford, R.J., 2012. Exploratory Data Analysis and Data Display, in: Birks, H.J.B., Lotter, A.F., Juggins, S., Smol, J.P. (Eds.), Tracking Environmental Change Using Lake Sediments, Developments in Paleoenvironmental Research 5, Developments in Paleoenvironmental Research. Springer Netherlands, Dordrecht, pp. 123–141. doi:10.1007/978-94-007-2745-8

Kanner, L.C., Burns, S.J., Cheng, H., Edwards, R.L., Vuille, M., 2013. High-resolution variability of the South American summer monsoon over the last seven millennia: insights from a speleothem record from the central Peruvian Andes. Quat. Sci. Rev. 75, 1–10. doi:10.1016/j.quascirev.2013.05.008

Leng, M.J., Marshall, J.D., 2004. Palaeoclimate interpretation of stable isotope data from lake sediment archives. Quat. Sci. Rev. 23, 811–831. doi:10.1016/j.quascirev.2003.06.012

Marcott, S. a, Shakun, J.D., Clark, P.U., Mix, A.C., 2013. A reconstruction of regional and global temperature for the past 11,300 years. Science 339, 1198–1201. doi:10.1126/science.1228026

Markgraf, V., 1989. Palaeoclimates in Central and South America since 18,000 BP based on Pollen and lake-level records. Quat. Sci. Rev. 8, 1–24.

Massaferro, J., Larocque, I.T., 2013. Using a newly developed chironomid transfer function for reconstructing mean annual air temperature at Lake Potrok Aike , Patagonia , Argentina. Ecol. Indic. 24, 201–210. doi:10.1016/j.ecolind.2012.06.017

Massaferro, J., Larocque-Tobler, I., Brooks, S.J., Vandergoes, M., Dieffenbacher-Krall, A., Moreno, P., 2014. Quantifying climate change in Huelmo mire (Chile, Northwestern Patagonia) during the Last Glacial Termination using a newly developed chironomid-based temperature model. Palaeogeogr. Palaeoclimatol. Palaeoecol. 399, 214–224. doi:10.1016/j.palaeo.2014.01.013

Matthews-Bird, F., Gosling, W.D., Coe, A.L., Bush, M., Mayle, F.E., Axford, Y., Brooks, S.J., 2015. Environmental controls on the distribution and diversity of lentic Chironomidae (Insecta : Diptera) across an altitudinal gradient in tropical South America. Ecol. Evol. 1–22. doi:10.1002/ece3.1833

Mayewski, P.A., Rohling, E.E., Curt Stager, J., Karlén, W., Maasch, K.A., David Meeker, L., Meyerson, E. a., Gasse, F., van Kreveld, S., Holmgren, K., Lee-Thorp, J., Rosqvist, G., Rack, F., Staubwasser, M., Schneider, R.R., Steig, E.J., 2004. Holocene climate variability. Quat. Res. 62, 243–255. doi:10.1016/j.yqres.2004.07.001

Meyer, I., Wagner, S., 2008. The Little Ice Age in southern Patagonia : Comparison between paleoecological reconstructions and downscaled model output of a GCM simulation. PAGES news 16.

Mosblech, N. a. S., Bush, M.B., Gosling, W.D., Hodell, D., Thomas, L., van Calsteren, P., Correa-Metrio, A., Valencia, B.G., Curtis, J., van Woesik, R., 2012. North Atlantic forcing of Amazonian precipitation during the last ice age. Nat. Geosci. 5, 817–820. doi:10.1038/ngeo1588

O’Brien, S.R., Mayewski, P. a., Meeker, L.D., Meese, D.A., Twickler, M.S., Whitlow, S.I., 1995. Complexity of Holocene Climate as Reconstructed from a Greenland Ice Core. Science (80-.). 270, 1962–1964.

Olander, H., Korhola, a., Blom, T., Birks, H.J.B., 1999a. An expanded calibration model for inferring lakewater and air temperatures from fossil chironomid assemblages in northern Fennoscandia. The Holocene 9, 279–294. doi:10.1191/095968399677918040

Olander, H., Korhola, a., Blom, T., Birks, H.J.B., 1999b. An expanded calibration model for inferring lakewater and air temperatures from fossil chironomid assemblages in northern Fennoscandia. The Holocene 9, 279–294. doi:10.1191/095968399677918040

Oldfield, F., Steffen, W., 2014. Anthropogenic climate change and the nature of Earth System science. Anthr. Rev. 1, 70–75. doi:10.1177/2053019613514862

Pinder, L.C.V., 1986. Biology of freshwater Chironomidae. Annu. Rev. Entomol. 31, 1–23.

Polissar, P.J., Abbott, M.B., Wolfe, a P., Bezada, M., Rull, V., Bradley, R.S., 2006. Solar modulation of Little Ice Age climate in the tropical Andes. Proc. Natl. Acad. Sci. U. S. A. 103, 8937–42.

doi:10.1073/pnas.0603118103

Prat, N., Rieradevall, M., Acosta, R., Villamarín, C., M, G.D.I.F.E., 2011. Las Larvas de Chironomidae (Diptera) DE Los rios Altoandinos de Ecuador y Peru, Clave par la determinacion de los generos.

Quinlan, R., Smol, J., 2001. Setting minimum head capsule abundance and taxa deletion criteria in chironomid-based inference models. J. Paleolimnol. 26, 327–342.

Rees, A.B.H., Cwynar, L.C., Cranston, P.S., 2008. Midges (Chironomidae, Ceratopogonidae, Chaoboridae) as a temperature proxy: a training set from Tasmania, Australia. J. Paleolimnol. doi:10.1007/s10933-008-9222-6

Reuter, J., Stott, L., Khider, D., Sinha, A., Cheng, H., Edwards, R.L., 2009. A new perspective on the hydroclimate variability in northern South America during the Little Ice Age. Geophys. Res. Lett. 36, L21706. doi:10.1029/2009GL041051

Rieradevall, M., Brooks, S., 2001. An identification guide to subfossil Tanypodinae larvae (Insecta: Diptera: Chrironomidae) based on cephalic setation. J. Paleolimnol. 81–99.

Self, A.E., Brooks, S.J., Birks, H.J.B., Nazarova, L., Porinchu, D., Odland, A., Yang, H., Jones, V.J., 2011. The distribution and abundance of chironomids in high-latitude Eurasian lakes with respect to temperature and continentality: development and application of new chironomid-based climate-inference models in northern Russia. Quat. Sci. Rev. 30, 1122–1141. doi:10.1016/j.quascirev.2011.01.022

Telford, Birks, H.J.B., 2011a. A novel method for assessing the statistical significance of quantitative reconstructions inferred from biotic assemblages. Quat. Sci. Rev. 30, 1272–1278. doi:10.1016/j.quascirev.2011.03.002

Telford, Birks, H.J.B., 2011b. Effect of uneven sampling along an environmental gradient on transfer-function performance. J. Paleolimnol. 46, 99–106. doi:10.1007/s10933-011-9523-z

ter Braak, C.J.F., 1987. Ordination, in: Jongman, R.., ter Braak, C.J.., van Tongeren, O.F.R. (Eds.), Data Analysis in Community Ecology. Pudoc, Wageningen, The Netherlands, pp. 91–173.

ter Braak, C.J.F., Juggins, S., 1993. Weighted averaging partial least squares regression (WA-PLS): an improved method for reconstructing environmental variables from species assemblages. Hydrobiologia 269/70, 485–502.

ter Braak, C.J.F., Looman, C.W.., 1986. Weighted averaging, logisitc regression and the Gaussian response model. Vegetatio 65, 3–11.

Thompson, L.., Mosley-Thompson, E., Dansgaard, W., Grootes, P.., 1986. The Little Ice Age as Recorded in the Stratigraphy of the Tropical Quelccaya Ice Cap. Science, 234, 361–364.

Thompson, L.G., Mosley-Thompson, E., Brecher, H., Davis, M., León, B., Les, D., Lin, P.-N., Mashiotta, T., Mountain, K., 2006. Abrupt tropical climate change: past and present. Proc. Natl. Acad. Sci. U. S. A. 103, 10536–43. doi:10.1073/pnas.0603900103

Thompson, L.G., Mosley-Thompson, E., Davis, M.E., 1995. Late glacial stage and Holocene tropical ice core records from Huascaran, Peru. Science, 269, 46–50.

Thompson, L.G., Mosley-Thompson, E., Davis, M.E., Henderson, K.A., Brecher, H.H., Zagorodnov, V.S., Mashiotta, T.A., Lin, P.-N., Mikhalenko, V.N., Hardy, D.R., Beer, J., 2002. Kilimanjaro ice core records: evidence of Holocene climate change in tropical Africa. Science 298, 589–93. doi:10.1126/science.1073198

Trivinho-Strixino, S., 2011. Larvas de Chironomidae guia de Identificacao. Universidade Federale de Sao Carlos.

van Geel, B., Raspopov, O.M., Renssen, H., Plicht, J. Van Der, Dergachev, V.A., Meijer, H.A.J., 1999. The role of solar forcing upon climate change. Quat. Sci. Rev. 18, 331–338.

Velle, G., Brodersen, K.P., Birks, H.J.B., Willassen, E., 2010. Midges as quantitative temperature indicator species: Lessons for palaeoecology. The Holocene 20, 989–1002. doi:10.1177/0959683610365933

Velle, G., Brooks, S.J., Birks, H.J.B., Willassen, E., 2005. Chironomids as a tool for inferring Holocene climate: an assessment based on six sites in southern Scandinavia. Quat. Sci. Rev. 24, 1429–1462. doi:10.1016/j.quascirev.2004.10.010

frazer 2/25/2016 3:33 PM
Deleted: (80-.)

frazer 2/25/2016 3:32 PM
Deleted: (80-.)

frazer 2/25/2016 3:34 PM
Deleted: h

frazer 2/25/2016 3:33 PM
Deleted: Thompson, L.G., Mosley-Thompson, E., Davis, M.E., Lin, P.N., Henderson, K.A., Cole-Dai, J., Bolzan, J.F., Liu, K.B., 1995. Late glacial stage and holocene tropical ice core records from huascaran, peru. Science 269, 46–50. doi:10.1126/science.269.5220.46

Vuille, M., Bradley, R., Keimig, F., 2000. Climate variability in the Andes of Ecuador and its relation to tropical Pacific and Atlantic sea surface temperature anomalies. J. Clim. 13, 2520–2535.

Walker, I.R., Cwynar, L.C., 2006. Midges and palaeotemperature reconstruction—the North American experience. Quat. Sci. Rev. 25, 1911–1925. doi:10.1016/j.quascirev.2006.01.014

Walker, I.R., Mathews, R.W., 1987. Chironomids, lake trophic status and climate. Quat. Res. 28, 431–437.

Wanner, H., Beer, J., Bütikofer, J., Crowley, T.J., Cubasch, U., Flückiger, J., Goosse, H., Grosjean, M., Joos, F., Kaplan, J.O., Küttel, M., Müller, S.A., Prentice, I.C., Solomina, O., Stocker, T.F., Tarasov, P., Wagner, M., Widmann, M., 2008. Mid- to Late Holocene climate change: an overview. Quat. Sci. Rev. 27, 1791–1828. doi:10.1016/j.quascirev.2008.06.013

Wiederholm, T., 1983. Chironomid of the Holarctic region. Keys and diagnosis. Part 1. Larvae. Entomologica Scandinavica Supplement 19.

Woodward, C., Shulmeister, J., 2006. New Zealand chironomids as proxies for human-induced and natural environmental change: transfer functions for temperature and lake production (chlorophyll a). J. Paleolimnol. 36, 407–429.

Wu, J., Porinchu, D.F., Horn, S.P., Haberyan, K. a., 2014. The modern distribution of chironomid sub-fossils (Insecta: Diptera) in Costa Rica and the development of a regional chironomid-based temperature inference model. Hydrobiologia 742, 107–127. doi:10.1007/s10750-014-1970-x

frazer 2/25/2016 3:34 PM
Deleted: .

Table captions

Table 1

frazer 2/26/2016 3:47 PM
Deleted:

frazer 2/26/2016 3:40 PM
Formatted: Font:+Theme Body

frazer 2/26/2016 3:39 PM
Deleted: 1

Summary of the physical and chemical properties of the 59 calibration data set lakes including the total number of head capsules retrieved from each lake and the concentration of head capsules per gram of sediment. MAT= mean annual temperature, MAP= mean annual precipitation, LOI=loss-on-ignition (550°C).

Table 2

AMS radiocarbon dates used for the age-depth model of Lake Pindo. SUERC: lab code (from NERC Radiocarbon Facility, East Kilbride); BS: Bulk sediment; W: Wood; WA: weighted average.

Table 3

frazer 2/26/2016 3:39 PM
Deleted: 2

Results of detrended canonical correspondence analysis (DCCA) using single constraining variables. MAT= mean annual temperature, WT=water temperature, MAP=mean annual precipitation, LOI= Loss-on-ignition.

Table 4

frazer 2/26/2016 3:40 PM
Deleted: 3

Summary of the performance statistics of chironomid-based MAT(°C) inference models developed using classical and Bayesian methods based on leave one out cross validation. Weighted averaging inverse and classical (WAinv, WAcla), Weighted averaging partial least squares (WA-PLS), coefficient of determinant between predicted and observed (r^2_{jack}), root mean squared error of prediction (RMSEP$_{jack}$) as % of the gradient.

frazer 2/26/2016 3:47 PM
Deleted:

frazer 2/26/2016 3:47 PM
Formatted: Indent: Left: 0", First line: 0"

Tables

Table 1

	Calibration data set				
	Minimum	Mean	Median	Maximum	Std dev
Conductivity (µs)	5.9	363	185	3205	579
Depth (m)	0.1	5	2.2	25	5.4
Elevation (m a.s.l)	150	3142	3845	4655	1459
Latitude (S)	0.1	11.2	14.2	17.3	6.2
Longitude (W)	64.4	71.6	70.3	78.4	4.5
LOI (%)	0	19	13	80	16
MAT (°C)	0.8	12	10	25	7
MAP (mm/year)	468	1222	769	4421	952
pH	5.7	8	7.9	10.2	1.1
Total Head Capsules	23	77	76	164	35
Water Temperature (°C)	5	15	13	33	6
Head capsule/gram	2	27	22	105	22

Table 2

Sample code	Depth (cm)	Sample type	Age (yr ^{14}C BP)	Age (cal yrs BP) 2σ	Age (cal yr BP) estimation (WA)
SUERC- 54395	46	W	334 ± 42	289-470	373
SUERC-47634	117	W	974 ± 36	769-923	835
SUERC-47635	245	W	1973 ± 39	1812-1943	1868
SUERC-47569	329	W	2335 ± 37	2293-2361	2279
SUERC-47572	410	W	2829 ± 39	2781-2991	2916
SUERC-48854	461	BS	3974 ± 45	4241-4447	4336

Table 3

Variable	Variance Explained (%)	λ_1/λ_2	*P*
MAT	12.93	1.431	0.001
MAP	10.3	0.900	0.001
WT	11.21	1.230	0.001
pH	6.23	0.500	0.001
LOI	3.23	0.239	0.062
Depth	2.44	0.190	0.240
Conductivity	2.34	0.179	0.296

frazer 2/26/2016 3:47 PM
Deleted: ... [8]

frazer 2/26/2016 3:47 PM
Formatted: Centered

frazer 3/24/2016 10:04 AM
Comment [1]: Table added after reviewer suggestion

frazer 2/26/2016 3:48 PM
Formatted: Indent: Left: 0", First line: 0"

frazer 2/26/2016 3:41 PM
Deleted: 2

Table 4

Model	R^2_{Jack}	$RMSEP_{jack}$	Mean $bias_{jack}$	Max $bias_{jack}$	% change
WA (inv)	0.890	2.404	-0.017	4.665	-
WA (cla)	0.890	2.475	-0.035	4.279	-2.936
WA-TOL (inv)	0.851	2.831	-0.182	6.498	-
WA-TOL (cla)	0.852	2.951	-0.211	7.350	-4.263
WA-PLS (1)	0.889	2.431	0.094	4.891	-
WA-PLS (2)	0.890	2.412	0.109	3.982	0.766
WA-PLS (3)	0.869	2.617	0.096	5.558	-8.483
WA-PLS (4)	0.866	2.659	0.199	5.922	-1.592
WA-PLS (5)	0.875	2.568	0.213	6.201	3.409
Bayesian	0.909	2.373	0.598	3.158	

frazer 2/26/2016 3:41 PM
Deleted: 3

frazer 2/26/2016 3:47 PM
Deleted:

Figure captions

Figure 1

Location of the calibration data set lakes (black circles) and Laguna Pindo (white triangle).

Figure 2

Sediment description, radiocarbon dates (^{14}C age) and age-depth models of Laguna Pindo. Key colour for sediment descriptions: Black or dark brown = organic rich sediments (peat and clay respectively); White = dark sandy intervals; Greenish =greenish sandy clay, not compacted; Yellow = sediment gap (no sediment).

frazer 2/26/2016 3:43 PM
Formatted: Line spacing: double

Figure 3

frazer 2/26/2016 3:43 PM
Deleted: 2

Figure 2: Canonical correspondence analysis (CCA) of the calibration data set lakes and environmental variables with elevation and longitude removed after variance inflation analysis. MAP=mean annual precipitation, MAT=mean annual temperature, WT= water temperature, LOI=loss-on-ignition. Grey circles denote calibration lakes, dark grey triangles mark species. All species could not be labelled due to crowding; instead nine important taxa have been marked as examples.

Figure 4

frazer 2/26/2016 3:43 PM
Deleted: 3

Chironomid taxa in the modern calibration dataset lakes. Lakes are ordered (top to bottom) from cold to warm and chironomids are ordered by occurrence from cold to warm lakes. Only taxa present in three or more lakes are included. Dashed line shows a gap in calibration data set lakes between 16-20 °C of the MAT gradient. Detrended canonical correspondence analysis (DCCA) constrained by MAT shows the taxon turnover across the gradient. Head capsule concentration (hc/gram) is also included.

Figure 5

frazer 2/26/2016 3:43 PM
Deleted: 4

Weighted-average and Bayesian optima (*solid grey circles*) and tolerances (*thick lines*) of the 55-chironomid taxa included in the calibration dataset, MAT Range (*dashed lines*). Taxa are organised by WA temperature optima from cold to warm.

Figure 6

frazer 2/26/2016 3:43 PM
Deleted: 5

Model performance of the best performing classical method (WA) and Bayesian approach. A=weighted averaging method; B=Bayesian method. WA: R$^2_{jack}$= 0.890, RMSEP$_{jack}$ = 2.404°C, Mean bias$_{jack}$= -0.017°C, Max bias $_{jack}$=4.665°C. Bayesian: R$^2_{jack}$= 0.909, RMSEP$_{jack}$ =2.373°C, Mean bias$_{jack}$= 0.598°C, Max bias $_{jack}$= 3.158°C.

frazer 2/25/2016 3:37 PM
Formatted: Font:+Theme Body, 11 pt

frazer 2/26/2016 3:43 PM
Deleted:

Figure 7

frazer 2/26/2016 3:43 PM
Deleted: 6

Diagram of fossil chironomid assemblage of Laguna Pindo. Five significant zones were identified using optimal partitioning with a broken stick model. Detrended canonical correspondence analysis (DCCA) constrained by calibrated radiocarbon age shows taxon turnover through time. Only taxa with relative abundances greater than 5% are shown. SD=standard deviation, hc/gram= head capsules per gram of wet sediment.

Figure 8

frazer 2/26/2016 3:43 PM
Deleted: 7

Chironomid-inferred mean annual temperatures (MAT) at Laguna Pindo using the WA inverse (grey) and Bayesian (black) models. Sample specific errors for the WA model are obtained through bootstrapping 999 cycles. Errors of the Bayesian reconstruction are site-specific uncertainties.

frazer 3/24/2016 9:38 AM
Deleted: Key late-Holocene climate events are shaded in grey. LIA=the range of the earliest and latest date for the Little Ice Age in South America (Polissar *et al*., 2006). 3500-2500 global cooling event (Mayewski *et al*., 2004), note, however, the Laguna Pindo record only extends to 3000 cal yrs BP.

Figure 9

frazer 2/26/2016 3:43 PM
Deleted: 8

Distribution of Laguna Pindo fossil samples (black circles) included passively within a CCA of the calibration data set lakes (grey circles) constrained using the significant environmental variables. MAP= mean annual precipitation, MAT= mean annual temperature, WT= water temperature. The first and last fossil sample in the sedimentary sequence has been labelled (total sediment depth); there are no directional trends through time. Calibration lakes that lie at similar elevations as Laguna Pindo have been labelled.

frazer 2/26/2016 3:43 PM
Formatted: Justified, Indent: Left: 0", First line: 0", Line spacing: double

frazer 2/26/2016 3:43 PM
Deleted: ... [9]

Figure 10

frazer 2/26/2016 3:43 PM
Deleted: 9

(left to right): Chironomid-inferred WA classical MAT with sample specific errors generated using bootstrapping. Bayesian reconstruction with sample specific errors. Goodness-of-fit of the fossil assemblages to temperature, vertical dotted line indicates the 90th percentile of squared residual distances of modern samples to first axis in a CCA; samples to the right of the line have a poor fit-to-temperature. Nearest modern analogue analysis, vertical dotted line indicates the 5th percentile of squared chord distances of the fossil samples in the modern calibration data set; samples to the right of the line have no good modern analogues. Detrended canonical correspondence analysis (DCCA) sample scores with radiocarbon age used as the sole constraining variable. Head capsule concentration per gram of sediment. Zones are derived from optimal partitioning of fossil assemblages using a broken stick model to define significant zones. Sq res dis= square residual distance; Sq chrd dis= square chord distance; SD units= standard deviation units; hc/gram=head capsule per gram of sediment.

Figure 11

frazer 2/26/2016 3:43 PM
Deleted: 0

Histogram of the proportion of variance in the chironomid MAT transfer function explained by 999 transfer functions trained on random environmental variables. Solid black line denotes the proportion of variance explained by the chironomid WA inverse MAT transfer function. Black dashed line marks the proportion of variance explained by the first axis of PCA of the fossil data. Grey dashed line marks the 95% variance of the random reconstructions.

Figure 12

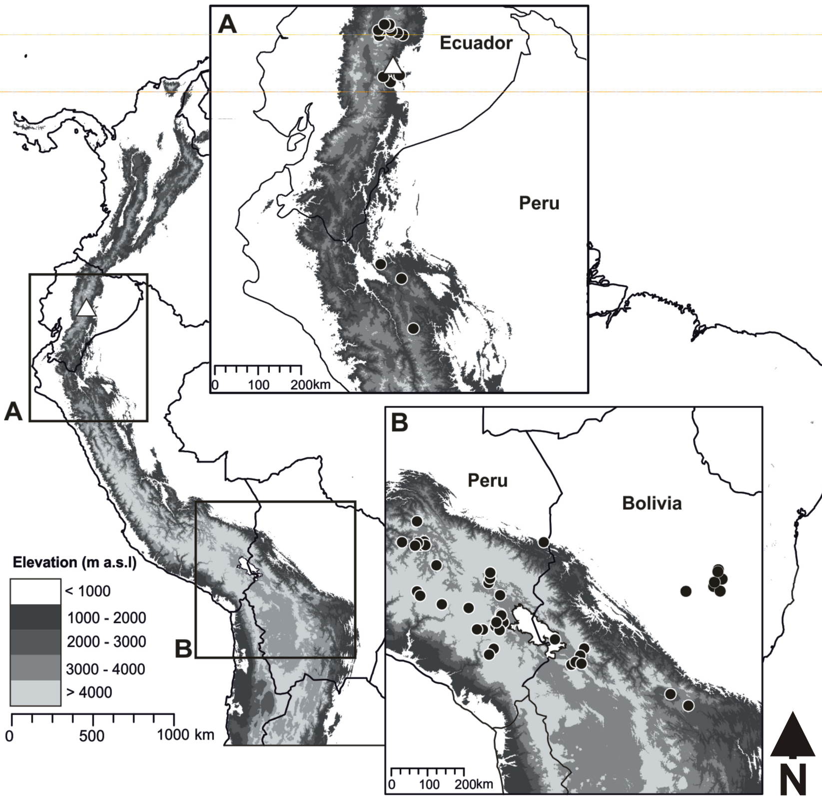
frazer 2/26/2016 3:43 PM
Deleted: 1

Individual likelihood functions for the fossil taxa in the coldest sample of the Laguna Pindo sequence (396 cm total depth, *c.* 2700 cal yr BP). The posterior probability distribution for temperature for the fossil sample is plotted in red, note this is plotted on an independent axis.

frazer 2/26/2016 3:43 PM
Deleted: ... [10]

Figures

Figure 1



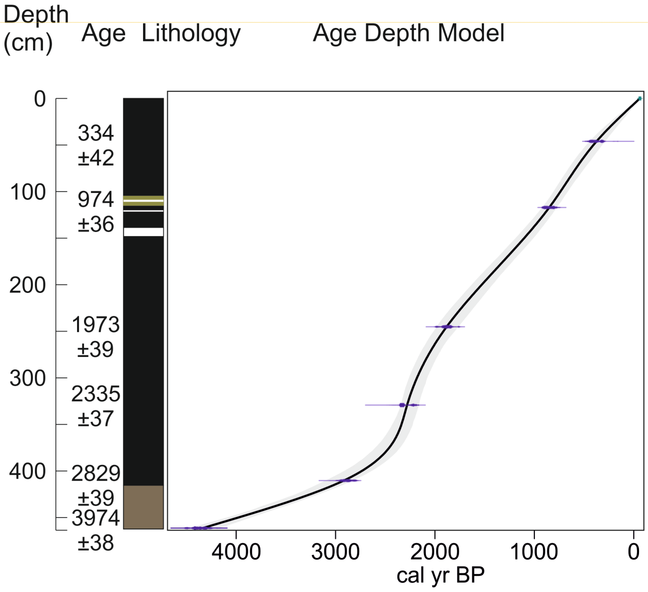
frazer 3/24/2016 10:05 AM
Comment [2]: North arrow added

Unknown
Formatted: Font:Calibri, Bold

frazer 2/26/2016 2:14 PM
Deleted: <sp>

Unknown
Formatted: Font:Bold

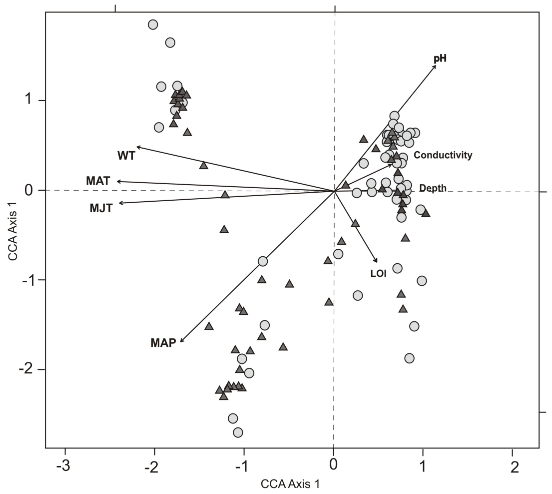
Figure 2



frazer 3/24/2016 10:05 AM
Comment [3]: Figure removed from supplementary and included in the manuscript

Unknown
Formatted: Font:Bold

Figure 3



frazer 3/24/2016 10:06 AM
Comment [4]: Latitude removed from the analysis

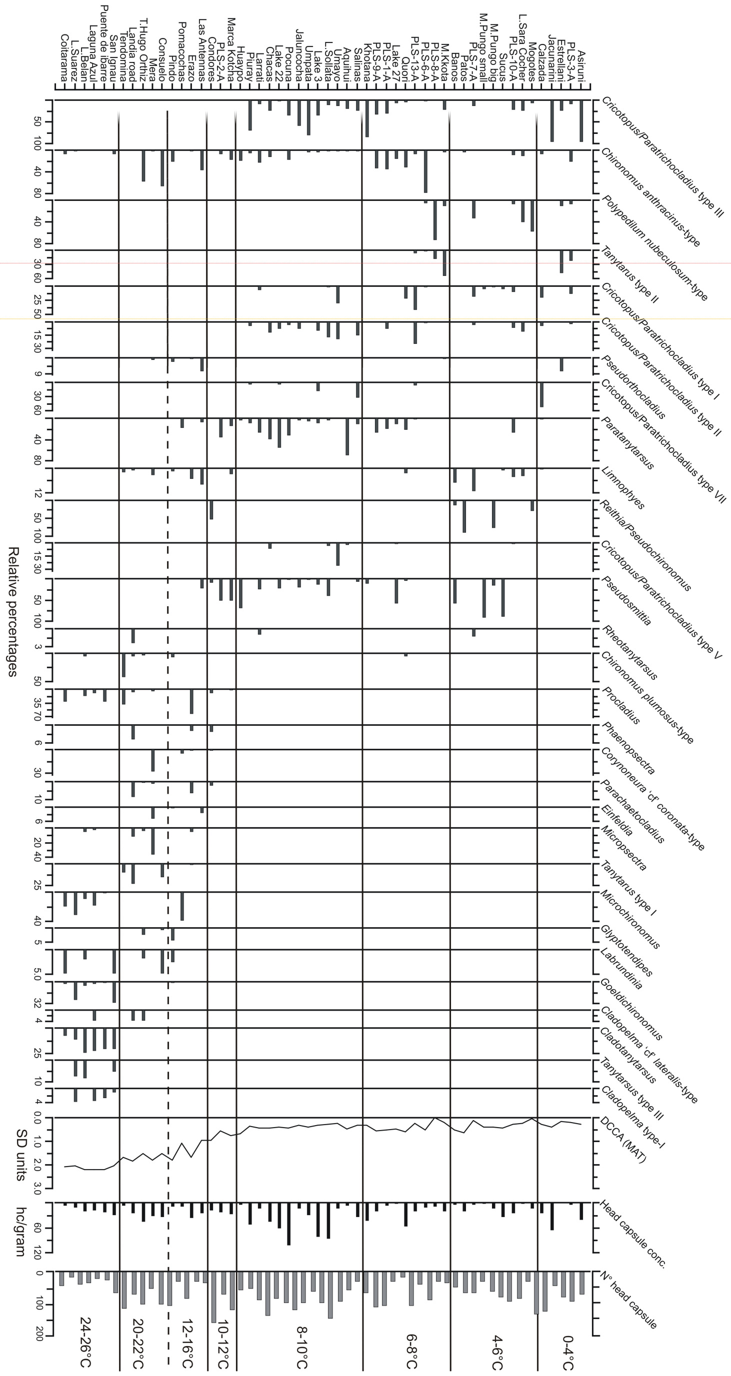
frazer 2/26/2016 3:44 PM
Deleted: 2

frazer 2/26/2016 2:15 PM
Deleted: *<sp>*

Unknown
Formatted: Font:Calibri

Unknown
Formatted: Font:Bold

Figure 4



frazer 2/26/2016 3:44 PM
Deleted: 3

frazer 2/26/2016 2:15 PM
Deleted: *<sp>*

Unknown
Formatted: Font:Bold

Figure 5

frazer 2/26/2016 3:44 PM
Deleted: 4

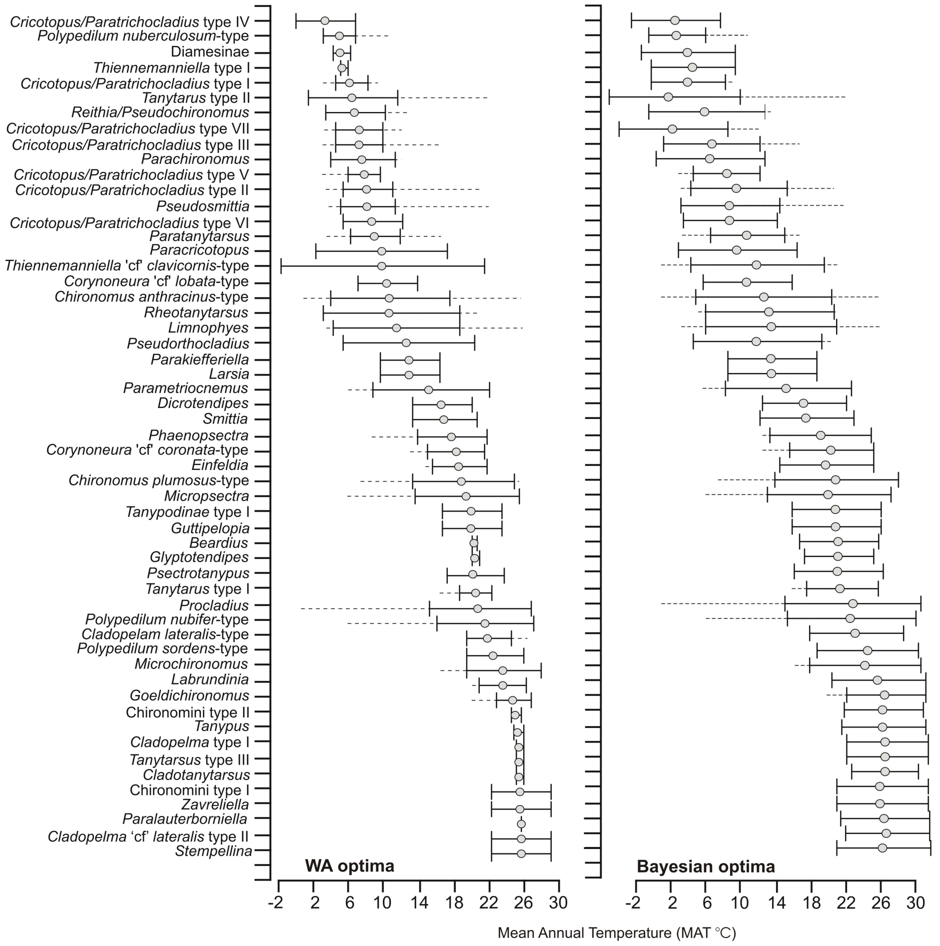


Figure 6

frazer 2/26/2016 3:44 PM
Deleted: 5

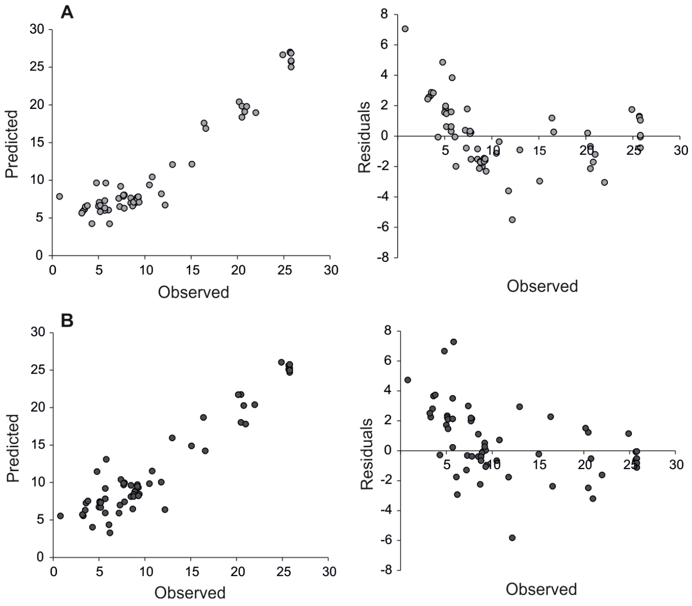
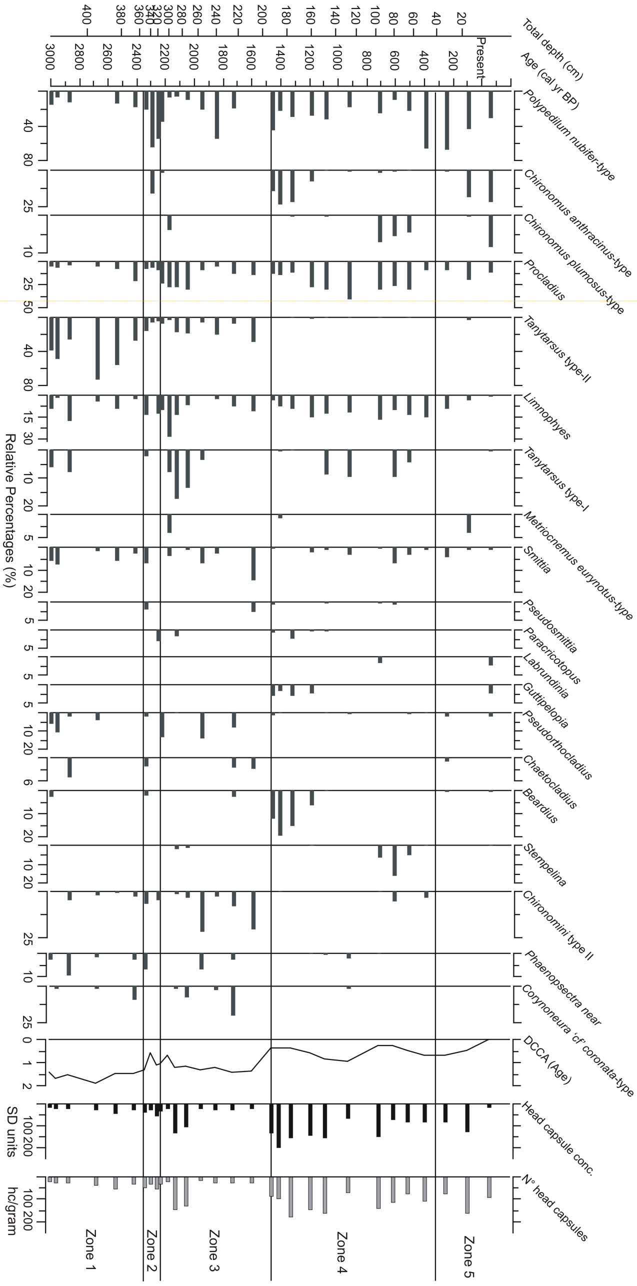


Figure 7



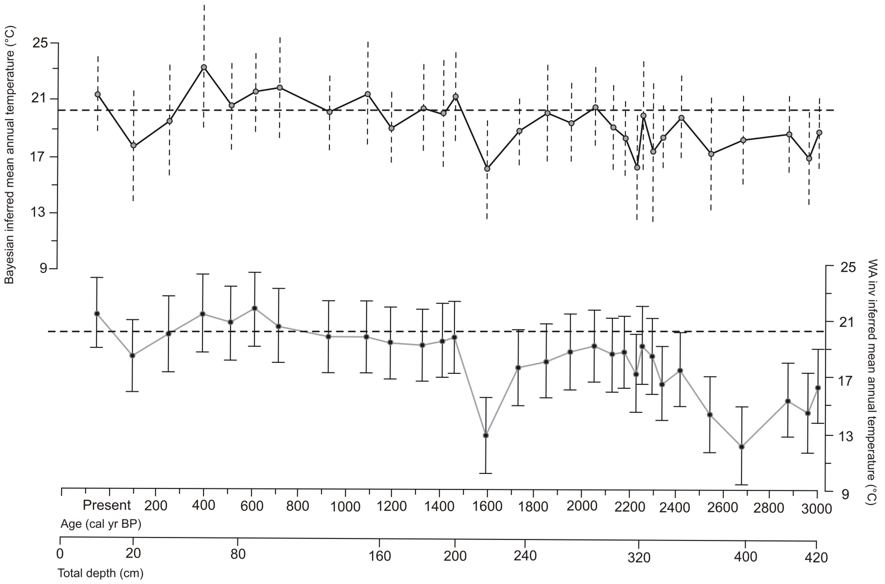
frazer 2/26/2016 3:44 PM

Deleted: 6

Unknown

Formatted: Font:Bold

Figure 8



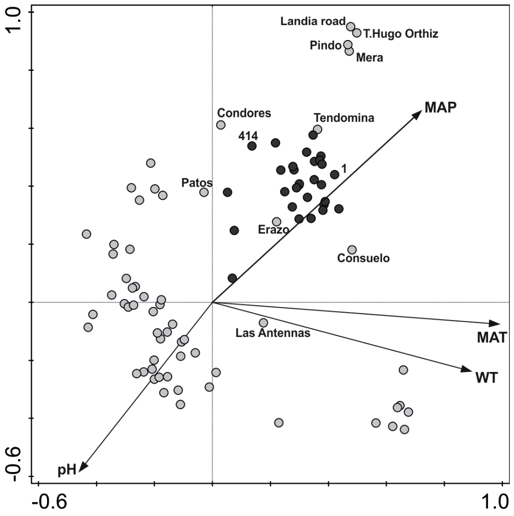
frazer 2/26/2016 3:44 PM

Deleted: 7

Unknown

Formatted: Font:Bold

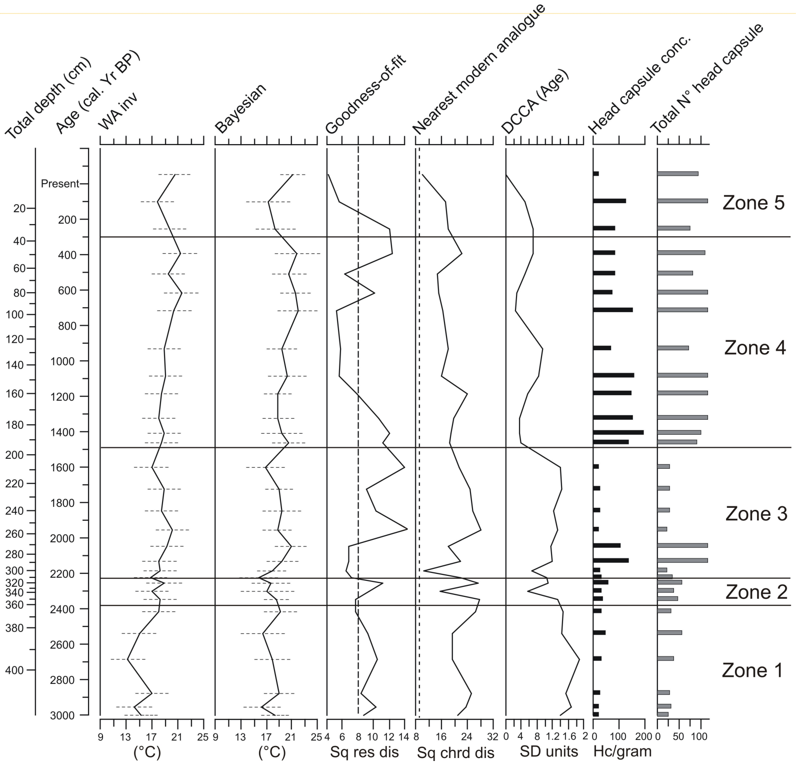
Figure 9



frazer 2/26/2016 3:44 PM
Deleted: 8

frazer 2/25/2016 2:42 PM
Deleted: ... [11]

Figure 10



frazer 2/26/2016 2:17 PM
Deleted: *<sp>*

Unknown
Formatted: Font:Bold

frazer 2/26/2016 3:44 PM
Deleted: 9

Unknown
Formatted: Font:Bold

Figure 11

frazer 2/26/2016 3:44 PM

Deleted: 0

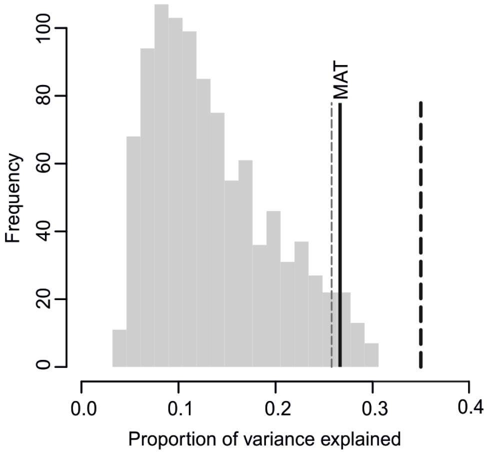


Figure 12

frazer 2/26/2016 3:45 PM
Deleted: 1