Title Page

Title

Seasonal lake-to-air temperature transfer functions derived from an analysis of 965 modern lakes: A tool for lacustrine proxy model comparison

Authors and Affiliations

Alexa Terrazas[*1,5], Nathan Hwangbo[2], Alexandrea J. Arnold[1,5], Robert N. Ulrich[3,5], Aradhna Tripati[*1,3,4,5,6]

[1]Department of Atmospheric and Oceanic Sciences, University of California, Los Angeles, USA

[2] Department of Statistics, University of California, Los Angeles, USA

[3] Department of Earth, Planetary, and Space Sciences, University of California, Los Angeles, USA

[4] Institute of the Environment and Sustainability, University of California, Los Angeles, USA

[5]Center for Diverse Leadership in Science, University of California, Los Angeles, USA

[6]School of Earth Science, University of Bristol, Bristol, UK

* Corresponding authors: aterrazas4@g.ucla.edu and atripati@g.ucla.edu

Running Head

Seasonal lake-to-air temperature transfer functions

Abstract

Lacustrine proxy reconstructions and paleoclimate modeling are important for understanding terrestrial climate change. Geochemical methods, including carbonate clumped isotope thermometry and TEX_{86}, have emerged as powerful tools for reconstructing past lake surface water temperature (LSWT). Nevertheless, global climate and earth system models often exclude lakes or simplify their representation due to limitations in grid resolution, computational intensity, variable complexity, and cost. These models prioritize the representation of large-scale atmospheric and oceanic processes that govern Earth's climate system, notably mean annual air temperature (MAAT). To facilitate proxy-model and proxy-proxy comparison, we develop new lake-to-air temperature transfer functions based on a comprehensive analysis of satellite and reanalysis data from 965 modern lakes. We use inverse modeling to examine LSWT during five seasonal intervals (spring, summer, warmest month, spring through summer, and annual) and assess the impact of climate variables, including latitude and elevation, on MAAT prediction from transfer functions. Compared to prior work, our models reduce random variations and biases, yielding more reliable predictions and improve generalizability, allowing for a broader range of conditions and locations like those of ancient lakes to be represented by our transfer functions. To showcase their utility, we employ the new transfer functions for reconstructing Pliocene-Pleistocene MAAT from lacustrine proxy data.

Main text

1. Introduction

Proxy-derived reconstructions of lake water temperature have been instrumental in advancing our understanding of terrestrial hydroclimate throughout the geologic record (Powers et al., 2005; Santi et al., 2020; Cheng et al., 2022). Lakes are optimal for studying climate change due to their sensitivity and rapid response to radiative forcing, and because of the widespread distribution and long-term preservation of lacustrine sediments (Adrian et al., 2009; Leavitt et al., 2009; Gierlowski-Kordesch, 2010). Lake-based proxy reconstructions are excellent paleoclimatic data sources, essential for assessing the efficacy of global climate models (Members, 1988; Braconnot et al., 2012; Oster et al., 2015). However, the disparity between directly simulated physical variables in models (e.g air temperature) and lake water temperature derived from paleoclimate proxies (e.g., Δ_{47}, $\delta^{18}O$, TEX_{86}) presents challenges for direct proxy-to-model comparisons, evaluating climate model performance, and validating proxy-based reconstructions.

Many paleoecological and paleoclimatological studies have utilized an inverse modeling approach to create transfer functions, empirical calibrations between proxy variables (e.g., tree ring width, pollen taxa, and microfossil assemblages) and physical (observed or simulated) variables (e.g., temperature and precipitation) (e.g. Imbrie and Kipp, 1971; Sachs et al., 1977; Bartlein et al., 1984; Birks, 1995). Previously, inverse modeling was utilized to develop seasonal lake-to-air temperature transfer functions, a method for translating proxy-derived reconstructions of lake surface water temperature (LSWT) to mean annual air temperature (MAAT) (Hren and

Sheldon, 2012). These lake-to-air temperature transfer functions have been widely used to interpret clumped isotope T(Δ_{47}) constraints on LSWT from lacustrine sediments (Huntington et al., 2015; Li et al., 2020; Santi et al., 2020; Cheng et al., 2022) as well as have been employed in proxy-model comparisons (Santi et al., 2020; Cheng et al., 2022). However, it is important to note the current calibrations were based on a dataset of 88 lakes with long-term temperature monitoring. This dataset is limited in monitoring location, duration, and geographic coverage, primarily representing lakes in northern latitudes with low elevations and mid-latitude regions, as revealed in the study by Hren and Sheldon (2012)

The recent development of modern satellite derived LSWT and air temperature reanalysis data products has presented an opportunity to conduct large-scale analyses capturing long-term global variability in lake water and air temperature across various climatological and hydrological settings (Sharma et al., 2015; Woolway et al., 2021; Wang et al., 2021). We hypothesize this advancement has the potential to reduce MAAT reconstruction uncertainties which is crucial when engaging in proxy-to-model comparisons. In this paper, we analyze data for 965 lakes to generate novel seasonal lake-to-air temperature transfer functions to facilitate paleoclimate reconstructions and proxy-model evaluation. We also examine the empirical link between LSWT and MAAT, by evaluating the importance of climate and geographical factors, such as latitude, elevation, total solar radiation, wind speed, cloud cover, and lake area. We hypothesize these previously unexplored variables might offer a more comprehensive explanation for the variance observed in the data and should be used as predictor variables in our transfer functions.

2. Methods

2.1 Lake and climatological data

We utilized long-term LSWT satellite-derived observation measurements from the Global Observatory of Lake Responses to Environmental Change (GloboLakes) project (Carrea and Merchant, 2019) and long-term monthly-averaged near-surface air temperature reanalysis data from ERA5-Land to capture the historical relationship between seasonal LSWT and MAAT (Muñoz Sabater, 2019). The Globolakes database contains 979 global lakes each with different LSWT time series, many spanning 1995 to 2016 at a spatial resolution of 0.025°. We obtain geographic and climatological information including lake elevation and surface area from the Global Lakes and Wetlands Database Level-1 (GLWD-1) (Lehner and Döll, 2004), monthly-averaged climatological reanalysis data from ERA5-Land (Muñoz Sabater, 2019), covering the period of 1950 through 2022 for 2-meter near-surface air temperature, 10m u-component of wind, and surface net solar radiation, at a resolution of 0.1°, and low cloud cover data at a resolution of 0.25° from ERA5 (Hersbach et al., 2023). We joined the separate datasets (Globolakes, GLWD-1, ERA5-Land, and ERA5) by geographical coordinates and lake identification codes (Supplemental Table 1). Lakes with missing elevations were excluded from the analysis. We compared whole lake mean LSWTs to lake center LSWTs, which is defined by GloboLakes as the point farthest from land. Given the similarity in results, we chose to use lake center LSWT in our analysis (Supplementary Fig. 1). However, the use of the whole lake average value does not change our conclusions.

2.2 Lacustrine proxy seasonality

Considering what is known of lacustrine proxy seasonality, our analysis concentrates on five intervals: (1) boreal April, May, June—equivalent to austral October, November, December; (2) boreal June, July, August—equivalent to austral December, January, February; (3) boreal April-October—equivalent to austral October-April; (4) warmest month; and (5) mean annual. These intervals are hereafter referred to as spring, summer, spring through summer, warmest month, and annual, respectively.

The impact of seasonal temperature biases on material-specific and individual species of carbonates for Δ_{47} and $\delta^{18}O$ thermometry and isoGDGT productivity for TEX_{86} has been a subject of investigation. Biogenic carbonates, including mollusks, are tolerant to narrow water temperature ranges (Dettman et al., 1999; Yan et al., 2009) and rely on peak photosynthetic activity and food availability for shell formation, which is most favorable during the spring through summer or the spring through early fall at mid-high northern latitude lakes (Versteegh et al., 2010; Apolinarska et al., 2015). Temperature reconstructions from abiotic and biologically mediated carbonates, such as micrites, marls, tufas, and microbialites, are often biased towards the warmest period of the year (Kano et al., 2003). This bias occurs due to enhanced evaporation, high microbial-induced photosynthesis reducing pCO_2, and increased lake water pH, which promotes high carbonate saturation (Platt & Wright, 2009). Significant correlations have been observed between TEX_{86} and mean annual LSWT derived from isoGDGTs (Powers et al., 2010, Loomis et al., 2017), indicating that isoGDGTs accumulate throughout the year. Nevertheless, recent research has highlighted that the productivity of archaea can be influenced by specific lake

properties, including mixing regime, size, water depth, and other environmental gradients (Urbach et al., 2001, Varela et al., 2008).

2.3 Statistical analysis

We investigated the effects of seasonal LSWT and climate on MAAT by statistical analysis. First, we analyzed the relationships among MAAT, seasonal LSWT, latitude, elevation, net solar radiation, wind speed, cloud cover, and lake area through bivariate correlation analysis as indicated by Pearson correlation coefficients. Recognizing the interdependence among climate variables, we analyzed loadings from principal component analysis (PCA) and calculated variance inflation factors (VIF) to evaluate the relative importance of each variable on principal component one and two and to detect multicollinearity. This approach enabled us to select features most appropriate for our analysis.

Following this, we proceeded to create a basic simple linear regression (SLR) model using ordinary least squares to investigate the relationship between seasonal LSWT and MAAT. The SLR model was then systematically refined through an iterative process by introducing additional predictor variables, chosen from on our preliminary assessments of correlation and feature importance. We evaluated each model generated through this approach using a comprehensive set of evaluation metrics, including the coefficient of determination (R^2), root mean squared error (RMSE), and Akaike information criterion (AIC), and examined the residuals and checked for deviations from linearity. To rigorously validate our models, we implemented k-folds cross-validation (Cooil et al., 1987), separating the dataset into an 90% in-sample training set and a 10% out-of-sample testing set.

3. Results

3.1 Geographical and seasonal temperature distribution of modern lakes

Our dataset includes 965 modern lakes, spanning a wide range of latitudes and elevations from 82°N to 55°S and -400 to 5400 meters above sea level, respectively (Fig. 1a). The size and distribution of lakes is largely shaped by geographical, climatic, and geological factors. While the majority (825) are in the Northern Hemisphere, a substantial portion (140) is also found in the Southern Hemisphere, and approximately two-thirds of all these lakes are situated at elevations ranging from -410 to 500 meters and surface areas from 48 up to 82000 km^2. Both the zonal and temperature-elevation profiles of LSWT demonstrate follow consistent patterns across each season, with varying rates of change (Fig. 1b, 1c). On average, seasonal LSWT decreases by 0.29 to 0.38°C per absolute degree latitude and 2.4 to 3.1°C per kilometer, while MAAT decreases by 0.54 °C per absolute degree latitude and 4.1 °C per kilometer. Based on the distribution of the data (Supplementary Figure 2), the lakes were subdivided into three latitude bins and five elevation bins. Seasonal LSWT consistently exceeds MAAT for low (0° to 30°), mid (30° to 60°), and high (60° to 90°) latitude lakes, as well as for each elevation category (Fig. 2). The largest disparity between seasonal LSWT and MAAT corresponds to high latitude and highly elevated (> 3000 meters above sea level) lakes, with differences of 10 to 20°C and 9 to 16°C, respectively, consistent with prolonged ice cover. In contrast, tropical lakes exhibit the smallest temperature differences, 3 to 5°C, which can be attributed to relatively stable and warm climatic conditions of low latitudes (0° to 30°) resulting in fewer pronounced seasonal changes. We also note a 7 to 15°C difference in seasonal LSWT and MAAT across different latitude and elevation categories for each season.

3.2 Identifying important climate variables

We quantitatively assessed the strength and direction of the relationship between pairwise combinations of seasonal LSWT and MAAT and possible governing variables (climate and lake characteristics) by correlation analysis. The heatmap in Fig. 3a shows the correlation structure between MAAT, seasonal LSWT, latitude, elevation, net solar radiation, wind speed, cloud cover, and lake area. Hypothesis testing reveals a strong ($r = 0.74$ to 0.96) statistically significant relationship between seasonal LSWT, latitude, and net solar radiation, and the target variable, MAAT ($p < 0.05$). Additionally, statistically significant relationships were detected among wind speed, low cloud cover, and elevation, albeit with lower correlation coefficients (ranging from 0.15 to 0.61). The loadings from the principal component analysis (PCA) shown in Fig. 3b reveal how climate variables influence the first two principal components, which capture the key patterns of data variability and provide valuable insights into the factors impacting predictions of MAAT. Principal component 1 (PC 1) shows strong associations with radiation, latitude, and LSWT (loadings > 0.45), while principal component 2 (PC 2) is primarily influenced by elevation and wind speed (loadings > 0.5). Considering the high correlation between radiation and latitude ($r = 0.94$; Fig. 3a) and the lack of robust proxy constraints, we exclude radiation as a predictor variable in our models. Similarly, we exclude low cloud cover and wind speed due to limited proxy constraints and their relatively small influence on PC 1 and PC 2 compared to seasonal LSWT, latitude, and elevation. Despite the moderately strong indirect relationship between seasonal LSWT and latitude, variance inflation factors remain low (Figure 3e), and we retain both variables in our analysis in accordance with the specific limitations of our study.

3.3 Seasonal lake-to-air temperature transfer functions

To investigate the effects of seasonal LSWT and climate on MAAT in more detail, we developed four distinct regression models varying in complexity. The initial SLR analysis, which we refer to as transfer function 1, revealed a robust positive correlation between long-term seasonal LSWT and MAAT (Figure 4). Transfer function 1 aptly captures the variation in the data across all seasons, with R^2 ranging from 0.75 to 0.93, indicating a strong fit. However, we observe greater variability and a larger spread in the residuals, for the warmest month and summer seasons (Fig. 4c and 4d), which transfer function 1 does not fully explain. Like Hren and Sheldon (2012), we consider the potential for a nonlinear relationship and introduce a second-degree polynomial LSWT term to transfer function 1, to create transfer function 2. This more complex model yields an 8% average reduction RMSE and a 4% increase in R^2 compared to transfer function 1 for the warmer seasonal regressions, as presented in Table 1. The lower AIC values favor transfer function 2 as a better model for the data across all seasons and because we incorporate the same predictor variables (LSWT and $LSWT^2$) into transfer function 2 as Hren and Sheldon (2012) do into their models, we can directly compare them. For the seasons modeled by both studies (spring, summer, spring though summer, and annual), transfer function 2 outperforms Hren and Sheldon's (2012) transfer functions indicated by a 10% average increase in R^2 an average RMSE reduction of 0.7°C.

In Section 3.2, we identified that incorporating latitude and elevation as explanatory variables in the regression analysis may help address underlying patterns in the data. By integrating latitude into our regressions via multiple linear regression (MLR), we developed transfer function 3. This led to substantial enhancements in model performance across all seasons, with an average 6% increase in R^2 and a reduction of 0.7°C in RMSE compared to

transfer function 2. Subsequently, the addition of elevation in transfer function 4 further fine-tuned our model, resulting in an average RMSE reduction of 0.8°C and a 5% increase in R^2 relative to transfer function 2. We employ a rigorous K-fold cross-validation approach to comprehensively assess the efficacy of our seasonal and mean annual transfer functions. We illustrate the residual distribution with overlaid boxplots on violin plots for both the newly developed and existing transfer functions in Fig. 5. In addition, we show how the conventional metrics of model performance compare across transfer functions as they apply to the k-fold cross validation procedure (Fig. 6).

3.4 Reconstructing paleo-MAAT

Below, we showcase two paleoclimate applications of transfer functions 1 through 4 (Fig. 7 and 8). The first case comes from Santi et al., (2020), which present hydroclimate (lake temperature, air temperature, evaporation, and precipitation) reconstructions for the northwestern Great Basin by measuring (Δ_{47}-LSWT) for lacustrine carbonates from Lake Surprise. The second case comes from Cheng et al., (2022), which use Pliocene-Pleistocene lacustrine (Δ_{47}-LSWT) to reconstruct MAAT from the Tibetan Plateau with the goal of constraining past and future permafrost carbon storage. In the first case, we opt to use the spring through summer transfer function due to the seasonality of the proxy used, which was tufa. On average, transfer functions 1 through 4 predict a 5.3 ± 3.2°C MAAT anomaly between the LGM and modern, which is roughly 1.2°C colder than the anomaly predicted using Hren and Sheldon (2012)'s transfer function (Figure 7).

Next, we examine variations in MAAT in the Tibetan Plateau during different geological

periods (Pliocene, mid-Pliocene warm period, and Northern Hemisphere Pleistocene). Applying transfer functions 1, 2, and 3 to their data, assuming consistent latitudinal positions of the samples over time, on average, we predict a -5.5 ± 5.9°C MAAT anomaly for the Pliocene, a -8.3 ± 5.8°C MAAT anomaly for the mid-Pliocene warm period, a -13.2 ± 5.1°C MAAT anomaly for the Northern Hemisphere Pleistocene, and a -12.1 ± 5.3°C MAAT anomaly for the Pleistocene. These estimates are 1.3°C and 0.7°C cooler and 1°C and 0.8°C warmer than Hren and Sheldon, respectively. Looking at the lowess regression of our MAAT reconstruction in Fig.8, we note that post the iNHG, MAAT remained below the conservative temperature limit for permafrost formation.

4. Discussion

The LSWT-MAAT transfer functions presented in this study are stable, robust, and general which enable more accurate lake-based assessments of climate change that limnologists, climate modelers, and paleoclimatologists can use, including to validate proxy reconstructions, evaluate models, and make decisions regarding freshwater resources. Metrics of transfer function skill were similar in cross-validation testing, which indicates that our approach appropriately handles out of sample data (Fig 6). This global-scale analysis has identified latitude and elevation as important variables that should be considered in calibrations of LSWT to MAAT (Fig. 3). Accounting for the variations introduced by latitude and elevation helped reduce prediction errors, likely due to their influence on climate conditions, which impact energy exchange at the lake-atmosphere boundary. Variation in net radiation flux is largely determined by latitude, therefore different latitudes exhibit varying LSWT patterns and seasonal temperature cycles (Straskraba, 1980). Altitude can influence air-water temperature relationships via

differential lapse rates (Livingstone et al., 1999), temperature gradients, seasonal variations in temperature, and atmospheric stability (Rueda et al., 2007), further impacting calibrations of lake water temperature to air temperature.

Nevertheless, our calibrations are based on three key assumptions: (1) seasonal LSWT along with absolute latitude and elevation can reliably predict changes in near surface mean annual air temperature; (2) the relationship between LSWT and MAAT is uniform through time; (3) modern data are sufficient to interpret paleoclimate proxy derived LSWT data. Therefore, transfer functions 1 through 4 will perform best for a given proxy derived LSWT in the range of sampled temperatures. However, MAAT can also be predicted for lakes with LSWTs outside this range through extrapolation, but predictions will be more prone to uncertainties. To address these uncertainties and further enhance the accuracy of paleoclimate reconstructions, future research should explore additional variables, including lake depth, and quantify the impact of their inclusion or exclusion in cases where paleoclimate records lack this information.

4.1 Applications: Paleolimnology

We specifically focused on how our lake-air temperature transfer functions can be used as tools for assessing the efficacy of global climate models and validating proxy-based reconstructions by transforming proxy-derived LSWT into MAAT, which we show in Figures 7 and 8. To constrain MAAT using lacustrine LSWT proxies, first the seasonal timescale of carbonate formation must be known or assumed. This allows the application of the appropriate seasonal transfer function for interpreting lacustrine Δ_{47} and or TEX_{86} data. We recommend using transfer function 4 when paleolatitude and paleoelevation are known, and transfer function

3 when paleoelevation is unknown since these demonstrate the best performance. However, if both paleolatitude and paleoelevation inputs are unknown, transfer function 2 should be used.

4.1.1 LGM hydroclimate in the northwestern Great Basin

Variations in air temperature predictions have consequential implications for hydroclimate predictions for lake-based evaporation and precipitation. Notably, cooler air temperatures (by ~ 1.2°C) and a consequently larger temperature anomaly of ~5°C (Figure 7) would support a more pronounced reduction in lake evaporation, which Santi et al. (2020) concluded played a critical role in growing and maintaining Lake Surprise during the LGM and deglacial period. However, with a reduction in lake evaporation rate, lake precipitation would also experience a decrease given the dependence on lake evaporation rate in the isotope mass balance model used to estimate precipitation. The overall effect on lake size will depend on the extent of changes in lake evaporation and precipitation reductions during the LGM.

4.1.1 Pliocene-Pleistocene temperatures in the Tibetan Plateau

Permafrost formation and destabilization is also sensitive to changes to the climate that are not well constrained. Cheng et al., (2022) used transfer functions from Hren and Sheldon (2012) to reconstruct Pliocene-Pleistocene MAAT from the Tibetan Plateau, the largest alpine permafrost region on the Earth. This approach allowed them to constrain historical and projected changes in alpine permafrost distributions, permafrost carbon storage, and compare their findings with climate models. Although the magnitude of our estimates using transfer function 3 differ from Cheng et al., (2022)'s estimates by roughly 2°C on average post the start of the iNHG and roughly 0.5°C prior to the start of the iNHG, Figure 8 highlights that our reconstructions still

supports permafrost formation post the iNHG to 0.85 Ma into the Pleistocene and permafrost destabilization prior to the start of the iNGH, during the Pliocene assuming a 0°C conservative temperature limit for permafrost formation.

5. Conclusions

This study establishes robust and generalized transfer functions linking LSWT to MAAT. Combining global observational datasets with machine learning and theory, we are able to introduce a powerful tool relating paleoclimate proxies to environmental variables. Our study overcomes several limitations of prior work, including the incorporation of a more diverse and extensive dataset of 965 lakes that spans a wide range of latitudes and elevations, which was previously underrepresented. Incorporation of latitude and elevation as predictor variables in the regressions, allows for the variance in the data to be better explained by our transfer functions. These variables, previously not considered in similar models, are revealed to have a substantial impact on LSWT-MAAT relationships, reflecting their influence on climate conditions and energy exchange between lakes and the atmosphere, enhancing the interpretation of historical and geological climate conditions and improving climate prediction. Revisiting previous MAAT paleo-reconstructions for southwestern North America during the Last Glacial Maximum and the Tibetan Plateau in the Pliocene-Pleistocene, highlight how past estimates and can change or remain the same and those changes can have implications on interpretations and evaluation of hydroclimate parameters or permafrost persistence.

6. References:

Adrian, R. et al. (2009) 'Lakes as sentinels of climate change', Limnology and oceanography, 54(6), pp. 2283–2297.

Apolinarska, K., Pełechaty, M. and Noskowiak, D. (2015) 'Differences in stable isotope compositions of freshwater snails from surface sediments of two Polish shallow lakes', Limnologica, 53, pp. 95–105. Available at: https://doi.org/10.1016/j.limno.2015.06.003.

Bartlein, P. J., Webb, T., & Fleri, E. (1984). Holocene Climatic Change in the Northern Midwest: Pollen-Derived Estimates. Quaternary Research, 22(3), 361–374. https://doi.org/10.1016/0033-5894(84)90029-2

Birks, H.J.B. 1995. Quantitative palaeoenvironmental reconstructions. In: Maddy, D. & Brew, J.S. Statistical Modelling of Quaternary Science Data. Technical guide 5, Quaternary Research Association, Cambridge, pp.116-254.

Braconnot, P. et al. (2012) 'Evaluation of climate models using palaeoclimatic data', Nature Climate Change, 2(6), pp. 417–424. Available at: https://doi.org/10.1038/nclimate1456.

Carrea, L. and Merchant, C.J. (2019) 'GloboLakes: Lake Surface Water Temperature (LSWT) v4.0 (1995-2016)'. Centre for Environmental Data Analysis (CEDA). Available at: https://doi.org/10.5285/76A29C5B55204B66A40308FC2BA9CDB3.

Cheng, F. et al. (2022) 'Alpine permafrost could account for a quarter of thawed carbon based on Plio-Pleistocene paleoclimate analogue', Nature Communications, 13(1), p. 1329. Available at: https://doi.org/10.1038/s41467-022-29011-2.

Cooil, B., Winer, R. S., & Rados, D. L. (1987). Cross-Validation for Prediction. Journal of Marketing Research, 24(3), 271–279. https://doi.org/10.1177/002224378702400303

Dettman, D.L., Reische, A.K. and Lohmann, K.C. (1999) 'Controls on the stable isotope composition of seasonal growth bands in aragonitic fresh-water bivalves (unionidae)', Geochimica et Cosmochimica Acta, 63(7), pp. 1049–1057. Available at: https://doi.org/10.1016/S0016-7037(99)00020-4.

Gierlowski-Kordesch, E.H. (2010) 'Chapter 1 Lacustrine Carbonates', in A.M. Alonso-Zarza and L.H. Tanner (eds) Developments in Sedimentology. Elsevier (Carbonates in Continental Settings: Facies, Environments, and Processes), pp. 1–101. Available at: https://doi.org/10.1016/S0070-4571(09)06101-9.

Hersbach, H., Bell, B., Berrisford, P., Biavati, G., Horányi, A., Muñoz Sabater, J., Nicolas, J., Peubey, C., Radu, R., Rozum, I., Schepers, D., Simmons, A., Soci, C., Dee, D., Thépaut, J-N. (2023): ERA5 monthly averaged data on single levels from 1940 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS), DOI: 10.24381/cds.f17050d7.

Hren, M.T. and Sheldon, N.D. (2012) 'Temporal variations in lake water temperature: Paleoenvironmental implications of lake carbonate δ18O and temperature records', Earth and Planetary Science Letters, 337–338, pp. 77–84. Available at: https://doi.org/10.1016/j.epsl.2012.05.019.

Huntington, K.W. et al. (2015) 'High late Miocene–Pliocene elevation of the Zhada Basin, southwestern Tibetan Plateau, from carbonate clumped isotope thermometry', Geological

Society of America Bulletin, 127(1–2), pp. 181–199. Available at: https://doi.org/10.1130/B31000.1.

Imbrie, J., and N.G. Kipp, A new micropaleontological method for quantitative paleoclimatology: Application to a (1971), Late Pleistocene Caribbean core, in The Late Cenozoic Glacial Ages, edited by K. Turekian, pp. 71-181, Yale Univ. Press, New Haven, Conn.

Kano, A. et al. (2003) 'Origin of annual laminations in tufa deposits, southwest Japan', Palaeogeography, Palaeoclimatology, Palaeoecology, 191(2), pp. 243–262. Available at: https://doi.org/10.1016/0031-0182(02)00717-4.

Leavitt, P.R. et al. (2009) 'Paleolimnological evidence of the effects on lakes of energy and mass transfer from climate and humans', Limnology and Oceanography, 54(6part2), pp. 2330–2348. Available at: https://doi.org/10.4319/lo.2009.54.6_part_2.2330.

Lehner, B. and Döll, P. (2004): Development and validation of a global database of lakes, reservoirs and wetlands. Journal of Hydrology 296/1-4: 1-22.

Li, H. et al. (2020) 'Factors controlling the oxygen isotopic composition of lacustrine authigenic carbonates in Western China: implications for paleoclimate reconstructions', Scientific Reports, 10(1), p. 16370. Available at: https://doi.org/10.1038/s41598-020-73422-4.

Livingstone, D. M., Lotter, A. F., & Walkery, I. R. (1999). The Decrease in Summer Surface Water Temperature with Altitude in Swiss Alpine Lakes: A Comparison with Air Temperature Lapse Rates. Arctic, Antarctic, and Alpine Research, 31(4), 341–352.

https://doi.org/10.1080/15230430.1999.12003319

Loomis, S.E. et al. (2017) 'The tropical lapse rate steepened during the Last Glacial Maximum', Science Advances, 3(1), p. e1600815. Available at: https://doi.org/10.1126/sciadv.1600815.

Members, C. (1988) 'Climatic Changes of the Last 18,000 Years: Observations and Model Simulations', Science, 241(4869), pp. 1043–1052.

Muñoz Sabater, J. (2019): ERA5-Land hourly data from 1950 to present. Copernicus Climate Change Service (C3S) Climate Data Store (CDS). DOI: 10.24381/cds.e2161bac

Oster, J.L. et al. (2015) 'Steering of westerly storms over western North America at the Last Glacial Maximum', Nature Geoscience, 8(3), pp. 201–205. Available at: https://doi.org/10.1038/ngeo2365.

Platt, N. H., & Wright, V. P. (2009). Lacustrine Carbonates: Facies Models, Facies Distributions and Hydrocarbon Aspects. InLacustrine Facies Analysis (pp. 57–74). John Wiley & Sons, Ltd. https://doi.org/10.1002/9781444303919.ch3

Powers, L. et al. (2010) 'Applicability and calibration of the TEX86 paleothermometer in lakes', Organic Geochemistry, 41(4), pp. 404–413. Available at: https://doi.org/10.1016/j.orggeochem.2009.11.009.

Powers, L.A. et al. (2005) 'Large temperature variability in the southern African tropics since the Last Glacial Maximum', Geophysical Research Letters, 32(8). Available at: https://doi.org/10.1029/2004GL022014.

391 Rueda, F., E. Moreno-Ostos, and L. Cruz-Pizarro (2007), Spatial and temporal scales of transport
392 during the cooling phase of the ice-free period in a small high-mountain lake, Aquat. Sci.,
393 69, 115–128, doi:10.1007/s00027-006-0823-8.

394 Sachs, H. M., Webb, T., & Clark, D. R. (1977). Paleoecological Transfer Functions. Annual
395 Review of Earth and Planetary Sciences, 5(1), 159–178.
396 https://doi.org/10.1146/annurev.ea.05.050177.001111

397 Santi, L.M. et al. (2020) 'Clumped isotope constraints on changes in latest Pleistocene
398 hydroclimate in the northwestern Great Basin: Lake Surprise, California', GSA Bulletin,
399 132(11–12), pp. 2669–2683. Available at: https://doi.org/10.1130/B35484.1.

400 Sharma, S. et al. (2015) 'A global database of lake surface temperatures collected by in situ and
401 satellite methods from 1985–2009', Scientific Data, 2(1), p. 150008. Available at:
402 https://doi.org/10.1038/sdata.2015.8.

403 Straskraba, M. (1980), The effects of physical variables on freshwater production: Analyses
404 based on models, in The Functioning of Freshwater Ecosystems, edited by E. D. LeCren,
405 pp. 13–84, Cambridge Univ. Press, Cambridge, U. K.

406 Urbach, E., Vergin, K. L., Young, L., Morse, A., Larson, G. L., & Giovannoni, S. J. (2001).
407 Unusual bacterioplankton community structure in ultra-oligotrophic Crater Lake.
408 Limnology and Oceanography, 46(3), 557–572.
409 https://doi.org/10.4319/lo.2001.46.3.0557

410 Varela, M. M., Van Aken, H. M., Sintes, E., & Herndl, G. J. (2008). Latitudinal trends of

411 Crenarchaeota and Bacteria in the meso- and bathypelagic water masses of the Eastern
412 North Atlantic. Environmental Microbiology, 10(1), 110–124.
413 https://doi.org/10.1111/j.1462-2920.2007.01437.x

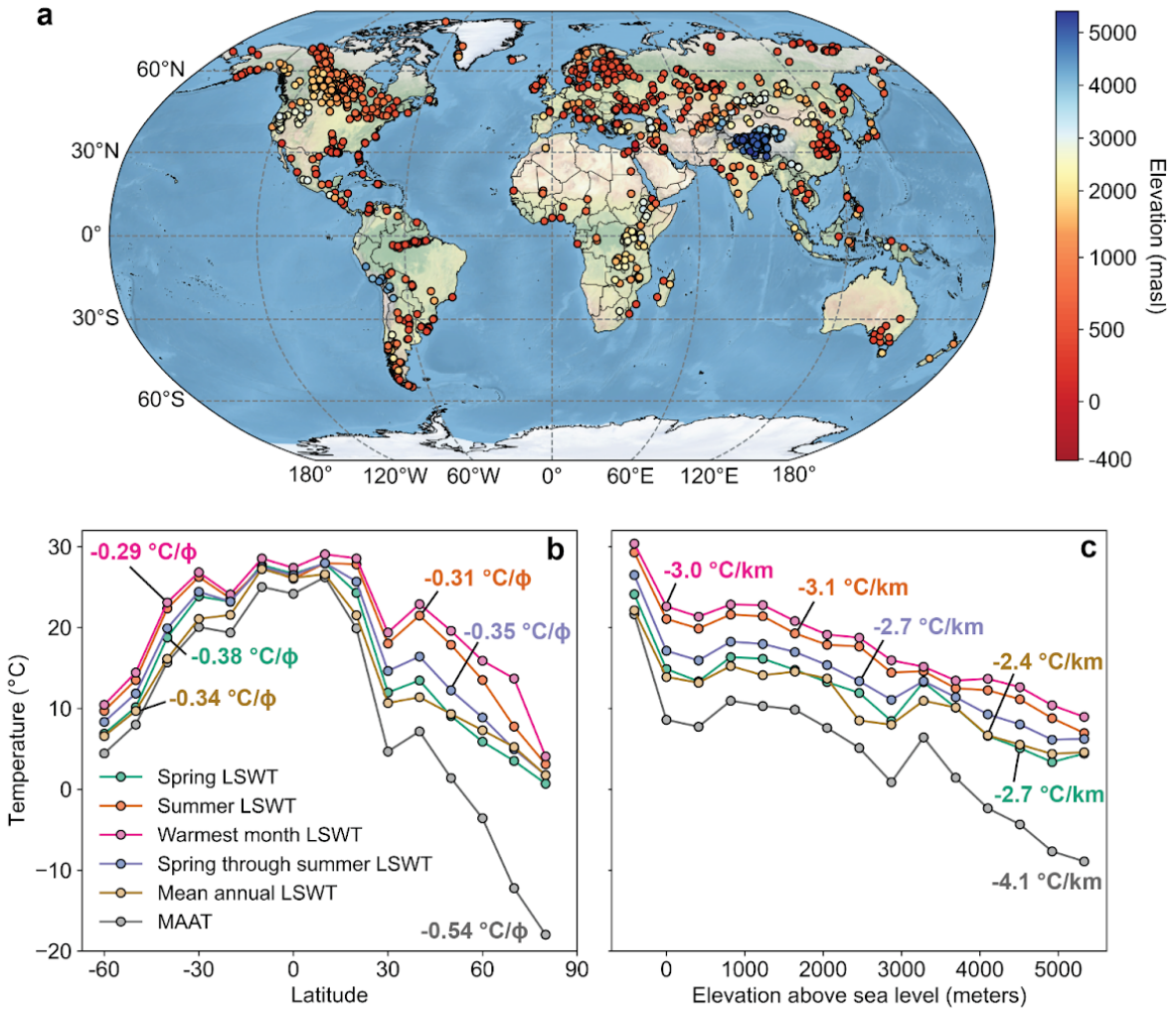
414 Versteegh, E.A.A. et al. (2010) 'Seasonally resolved growth of freshwater bivalves determined
415 by oxygen and carbon isotope shell chemistry', Geochemistry, Geophysics, Geosystems,
416 11(8). Available at: https://doi.org/10.1029/2009GC002961.

417 Wang, S., He, Y., Hu, S., Ji, F., Wang, B., Guan, X., & Piccolroaz, S. (2021). Enhanced
418 Warming in Global Dryland Lakes and Its Drivers. Remote Sensing, 14(1), 86.
419 https://doi.org/10.3390/rs14010086

420 Woolway, R.I., Jennings, E., Shatwell, T. et al. Lake heatwaves under climate change. Nature
421 589, 402–407 (2021). https://doi.org/10.1038/s41586-020-03119-1

422 Yan, H. et al. (2009) 'Stable isotope composition of the modern freshwater bivalve Corbicula
423 fluminea', Geochemical Journal, 43(5), pp. 379–387. Available at:
424 https://doi.org/10.2343/geochemj.1.0035.

425 *Figures*



426

427 Figure 1: Geographic distribution of lake localities, seasonal lake surface water temperature
428 (LSWT), and mean annual air temperature (MAAT). (a) Map of 965 modern lakes, characterized
429 by elevation relative to sea level. (b) Temperature versus latitude lapse rate presented as the
430 zonal mean across 10° latitude bands, where φ refers to absolute latitude in degrees. (c)
431 Temperature versus elevation lapse rate across 410 m elevation bins.

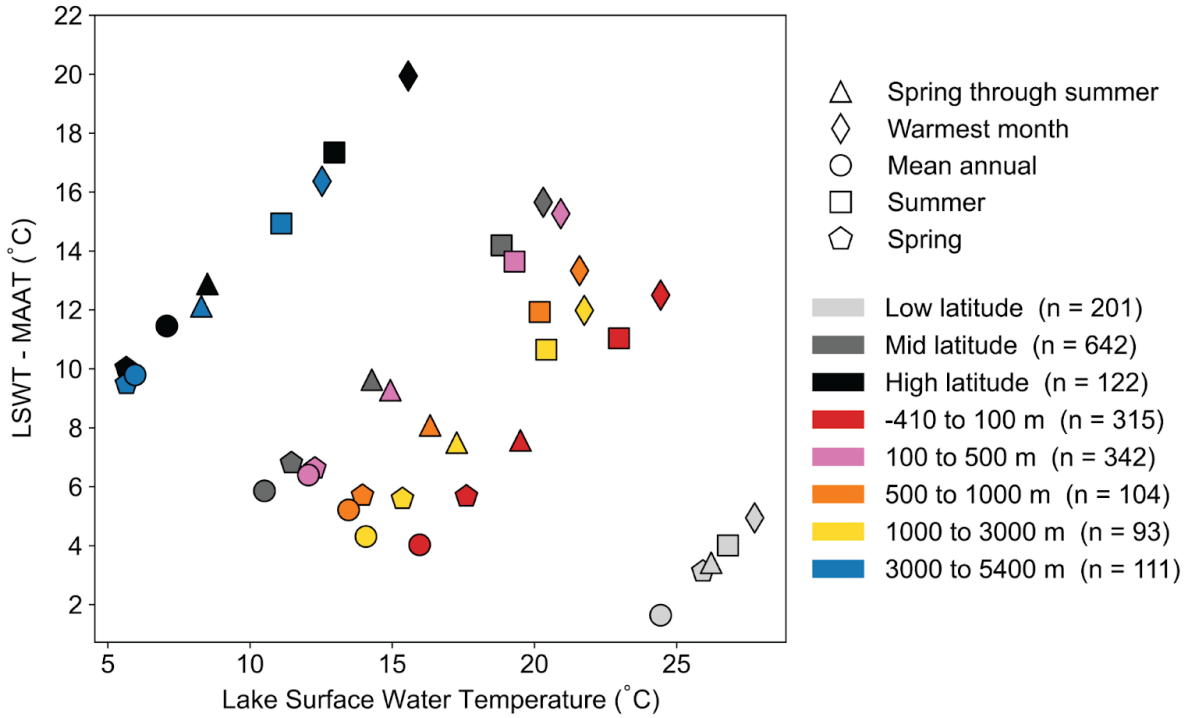
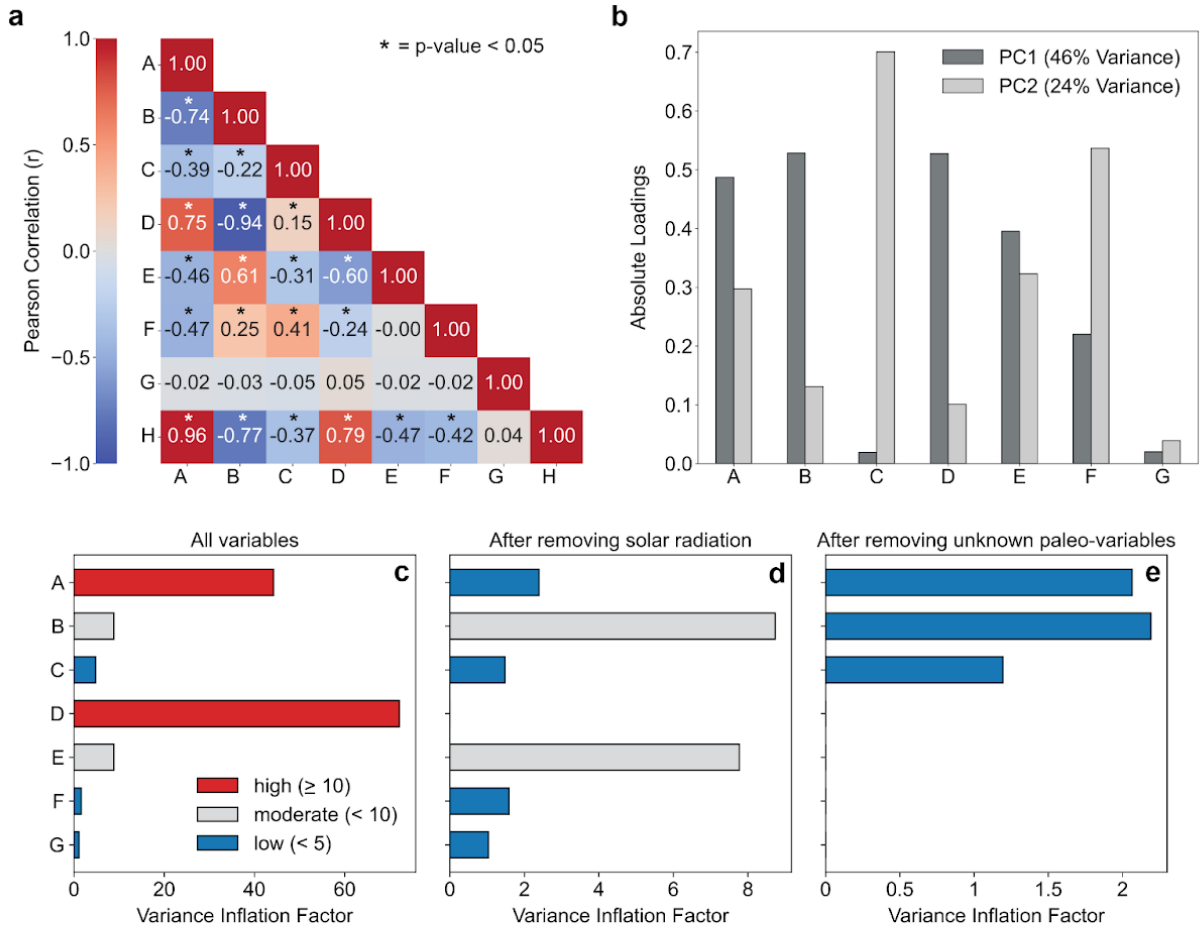
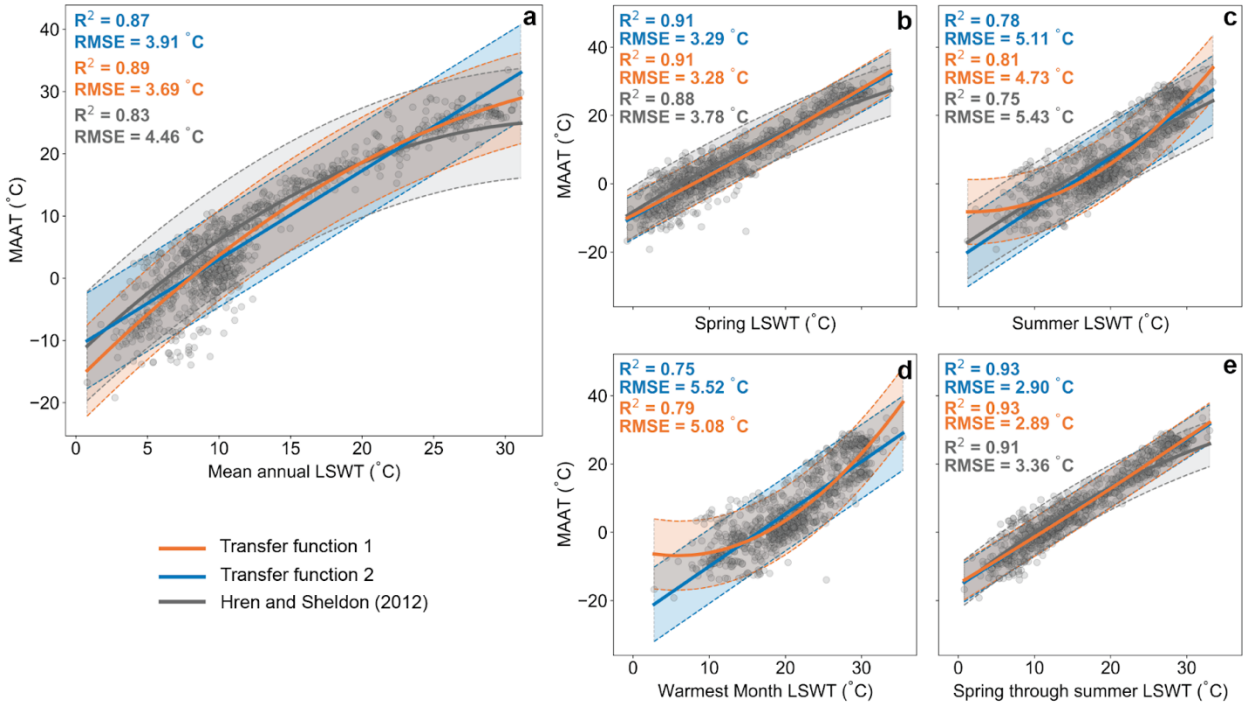


Figure 2: Average temperature differences in seasonal lake surface water temperature (LSWT) and mean annual air temperature (MAAT) for each latitude and elevation category. Latitude bins are defined into three categories: low (0° to 30°), mid (30° to 60°), and high (60° to 90°). Elevations bins are subdivided into five categories: -410 to 100 meters, 100 to 500 meters, 500 to 1000 meters, 1000 to 3000 meters, and 3000 to 5400 meters.



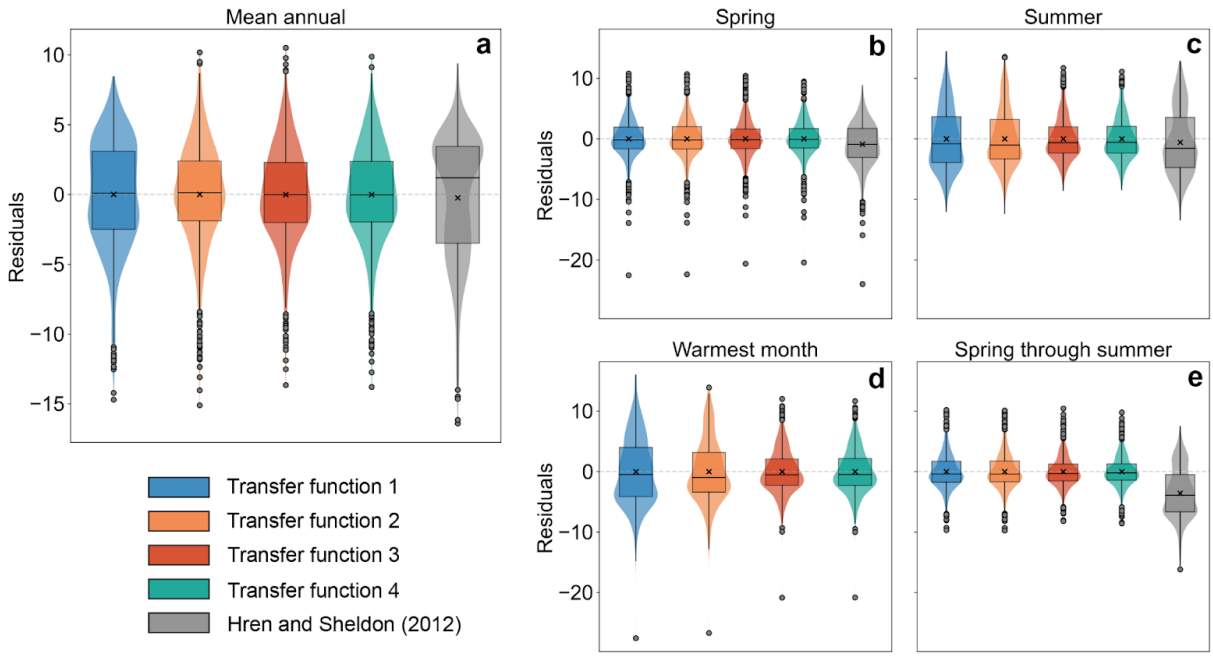
438

439 **Figure 3:** Lake and climate variables (a) The absolute loadings derived from the principal
440 component analysis (PCA) depict the individual contributions of each variable to principal
441 component 1 (PC 1) and principal component 2 (PC 2). (b) Pearson correlation heatmap matrix
442 with significance level expressed by asterisk (*p-value ≤ 0.05). The variables include: A =
443 Spring through summer lake surface water temperature, B = absolute latitude, C = lake elevation,
444 D = annual net solar radiation, E = annual low cloud cover, F = annual 10m horizontal
445 component of wind, G = lake surface area, H = mean annual air temperature.



446

447 **Figure 4:** Comparative analysis of transfer function 1 and 2. Plots (a) through (e) illustrate the
448 fitted curves and 95% prediction intervals for each model across different temporal segments: (a)
449 Spring, (b) Summer, (c) Warmest Month, (d) Spring through Summer, and (e) Mean Annual.
450 The efficacy of the models is evaluated through R-squared (R^2) and root mean square error
451 (RMSE) metrics.



452

453 **Figure 5**: Assessment of cross validated out of sample residual disparities between observed and

454 transfer function predicted mean annual air temperature (MAAT). Panels (a) through (e)

455 showcase the distribution of residuals, for each season: (a) mean annual, (b) spring, (c) summer,

456 (d) warmest month, and (e) spring through summer using transfer function 1 through 4.

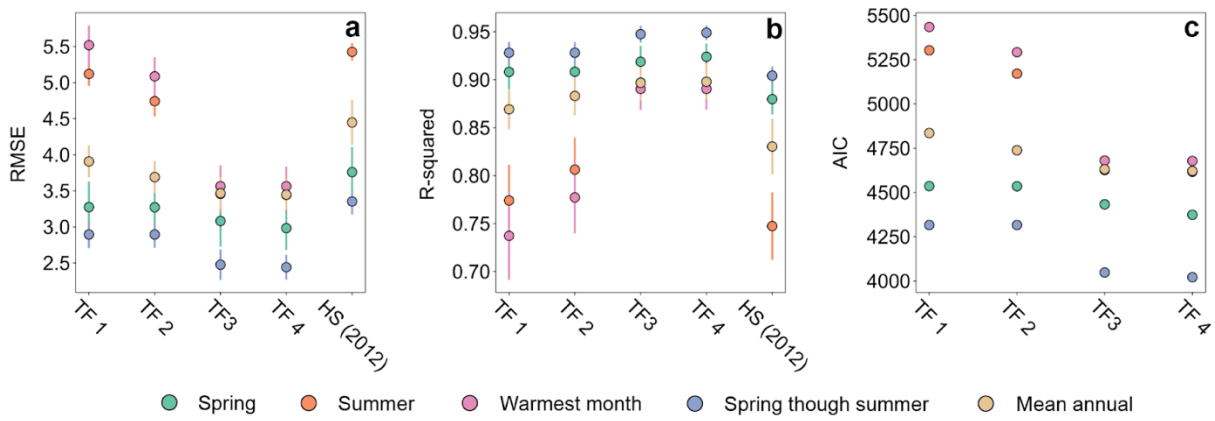


Figure 6: K-Fold cross-validated performance metrics comparison. (a) Root mean square error (RMSE) and (b) coefficient of determination (R^2) among transfer functions 1 through 4, and Hren and Sheldon (2012) across five proxy seasons: spring, summer, warmest month, spring through summer, and annual. The error bars depicted in the figures represent the variability in the performance metric at one standard deviation (1σ) level, derived from a comprehensive analysis spanning ten different folds in the K-Fold cross-validation procedure.

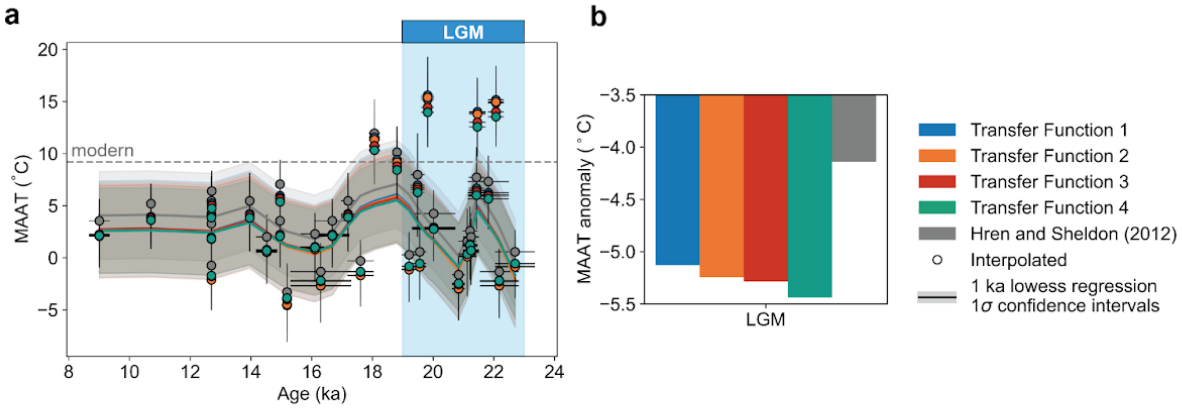
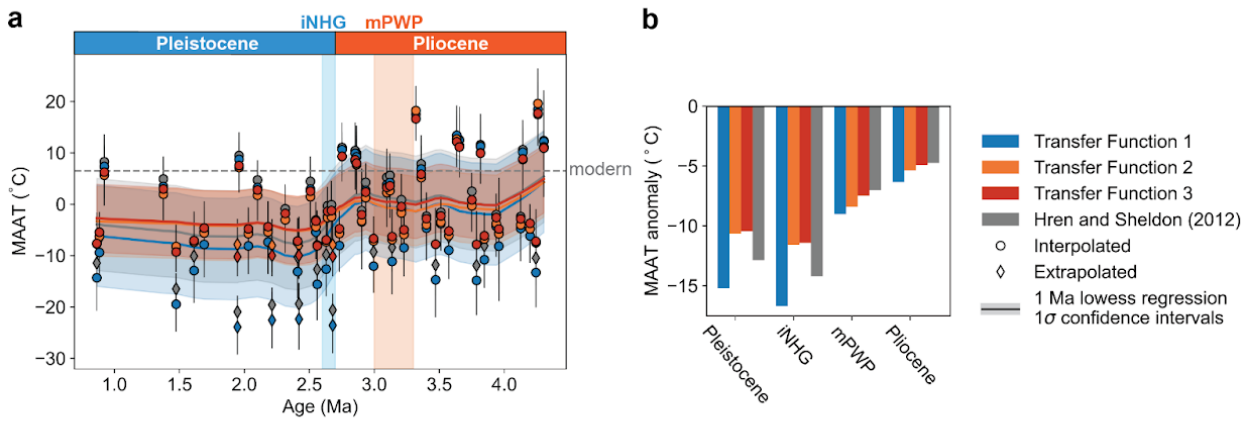


Figure 7: Reconstruction of Late Pleistocene hydroclimate in the Northwestern Great Basin (Lake Surprise). Calculations use published Δ_{47} lake temperatures (Santi et al., 2020) and transfer functions from this study and Hren and Sheldon (2012). Results shown are Last Glacial Maximum (LGM) mean annual air temperatures (MAAT), modern MAAT, and LGM to modern MAAT anomalies. (a) MAAT reconstruction against sample age. The reconstructions are based on transfer functions 1 (blue), 2 (red), 3 (red), and 4 (green) utilizing Δ_{47} proxy-derived lake surface water temperature (LSWT). The dashed line represents the locally weighted scatterplot smoothing (lowess) regression, which captures the underlying MAAT trends in the data by smoothing out local fluctuations. (b) LGM MAAT anomaly (LGM minus modern MAAT). The modern MAAT used is 9.1 ± 1°C (Santi et al., 2020). 1σ errors are calculated in quadrature.



475

476 **Figure 8:** Reconstruction of Pliocene-Pleistocene climate for the Tibetan Plateau (Kunlun Pass)
477 place. Calculations use published Δ_{47} lake temperatures (Cheng et al. 2022) and transfer
478 functions from this study or Hren and Sheldon (2012). Results shown are mean annual air
479 temperatures (MAAT), modern temperatures, and anomalies defined relative to modern MAAT
480 (data source). Intervals examined are the Pliocene, Pleistocene, mid-Pliocene warm period
481 (mPWP), and the intensification of the NHG (iNHG). (a) MAAT reconstruction against sample
482 age. The reconstructions are based on transfer functions 1 (blue), 2 (orange), and 3 (red), and
483 utilize Δ_{47} proxy-derived lake surface water temperature (LSWT). The dashed line represents the
484 locally weighted scatterplot smoothing (lowess) regression, which captures the underlying
485 MAAT trends by smoothing out local fluctuations. Diamonds indicate temperatures extrapolated
486 outside of calibration range. (b) Anomalies calculated using reconstructed temperatures relative
487 to modern MAAT. Modern temperature is 6.5±1°C. 1σ errors are calculated in quadrature.

488

489 *Tables*

Transfer Function	Predictor variables	Season	Equation	R^2	RMSE	AIC
This study: 1	LSWT	Spring	MAAT (°C) = 1.24 * LSWT - 9.74	0.91	3.29	5040
This study: 2	LSWT, $LSWT^2$	Spring	MAAT (°C) = 1.14 * LSWT + 0.0031 * $LSWT^2$ - 9.19	0.91	3.28	5039
This study: 3	LSWT, $LSWT^2$, φ	Spring	MAAT (°C) = 1.11 * LSWT + 0.0015 * $LSWT^2$ - 0.0472 * φ - 6.63	0.92	3.09	4925
This study: 4	LSWT, $LSWT^2$, φ, z	Spring	MAAT (°C) = 0.977 * LSWT + 0.0036 * $LSWT^2$ - 0.0597 * φ - 0.0006 * z - 4.41	0.93	2.99	4860
Hren and Sheldon (2012)	LSWT, $LSWT^2$	Spring	MAAT (°C) = -0.0097 * $LSWT^2$ + 1.379 * LSWT - 9.227	0.88	3.78	—
This study: 1	LSWT	Summer	MAAT (°C) = 1.47 * LSWT - 21.9	0.78	5.11	5892
This study: 2	LSWT, $LSWT^2$	Summer	MAAT (°C) = -0.139 * LSWT + 0.0418 * $LSWT^2$ - 8.11	0.81	4.73	5745
This study: 3	LSWT, $LSWT^2$, φ	Summer	MAAT (°C) = 0.301 * LSWT + 0.0256 * $LSWT^2$ - 0.126 * φ - 5.37	0.90	3.46	5140
This study: 4	LSWT, $LSWT^2$, φ, z	Summer	MAAT (°C) = 0.224 * LSWT + 0.0266 * $LSWT^2$ - 0.130 * φ - 0.0003 * z - 3.80	0.90	3.43	5130
Hren and Sheldon (2012)	LSWT, $LSWT^2$	Summer	MAAT (°C) = -0.0055 * $LSWT^2$ + 1.476 * LSWT - 18.915	0.75	5.43	—
This study: 1	LSWT	Warmest month	MAAT (°C) = 1.53 * LSWT - 25.3	0.75	5.52	6038
This study: 2	LSWT, $LSWT^2$	Warmest month	MAAT (°C) = -0.605 * LSWT + 0.0515 * $LSWT^2$ - 5.66	0.79	5.08	5880
This study: 3	LSWT, $LSWT^2$, φ	Warmest month	MAAT (°C) = 0.276 * LSWT + 0.0257 * $LSWT^2$ - 0.140 * φ - 6.25	0.90	3.56	5200
This study: 4	LSWT, $LSWT^2$, φ, z	Warmest month	MAAT (°C) = 0.202 * LSWT + 0.0268 * $LSWT^2$ - 0.142 * φ - 0.0002 * z - 4.97	0.90	3.56	5198
Hren and Sheldon (2012)	LSWT, $LSWT^2$	Warmest month	—	—	—	—
This study: 1	LSWT	Spring through summer	MAAT (°C) = 1.44 * LSWT - 15.8	0.93	2.90	4795
This study: 2	LSWT, $LSWT^2$	Spring through summer	MAAT (°C) = 1.34 * LSWT + 0.0027 * $LSWT^2$ - 15.1	0.93	2.89	4795
This study: 3	LSWT, $LSWT^2$, φ	Spring through summer	MAAT (°C) = 1.35 * LSWT - 0.0011 * $LSWT^2$ - 0.0623 * φ - 11.9	0.95	2.48	4498
This study: 4	LSWT, $LSWT^2$, φ, z	Spring through summer	MAAT (°C) = 1.25 * LSWT + 0.0007 * $LSWT^2$ - 0.0689 * φ - 0.0004 * z - 10.2	0.95	2.44	4468
Hren and Sheldon (2012)	LSWT, $LSWT^2$	Spring through summer	MAAT (°C) = -0.0146* $LSWT^2$ + 1.753 * LSWT - 16.079	0.91	3.36	—
This study: 1	LSWT	Mean annual	MAAT (°C) = 1.42 * LSWT - 11.2	0.87	3.91	5372
This study: 2	LSWT, $LSWT^2$	Mean annual	MAAT (°C) = 2.30 * LSWT - 0.0267 * $LSWT^2$ - 16.6	0.89	3.69	5265
This study: 3	LSWT, $LSWT^2$, φ	Mean annual	MAAT (°C) = 2.22 * LSWT - 0.0282 * $LSWT^2$ - 0.0544 * φ - 13.5	0.90	3.46	5145
This study: 4	LSWT, $LSWT^2$, φ, z	Mean annual	MAAT (°C) = 2.08 * LSWT - 0.0249 * $LSWT^2$ - 0.0610 * φ - 0.0003 * z - 11.8	0.90	3.44	5134
Hren and Sheldon (2012)	LSWT, $LSWT^2$	Mean annual	MAAT (°C) = -0.0318 * LSWT + 2.195 * LSWT - 12.607	0.83	4.46	—

490

491 **Table 1:** Evaluation of modeling approaches. Transfer function 1 and 2 are single variable
492 regressions dependent on seasonal LSWT. Transfer function 3 is a multiple linear regression
493 with seasonal LSWT, and latitude (φ) as predictors. Transfer function 4, is a multiple linear
494 regression with seasonal LSWT, latitude, and elevation (z) as predictors. Transfer functions are
495 evaluated with these performance metrics, coefficient of determination (R^2), root-mean-square
496 error (RMSE), mean error (ME), and Akaike information criterion (AIC). Model evaluation
497 refers to the population of all lakes, while model cross validation metrics apply only to k-fold
498 validation procedures.

Study Site	Time interval	Season	Transfer Function	Predictor variables	MAAT (°C)	MAAT (°C) error	MAAT (°C) anomaly	MAAT (°C) anomaly error
Lake Surprise	LGM	Spring through summer	This study: 1	LSWT	4.1	3.4	-5.1	3.4
Lake Surprise	LGM	Spring through summer	This study: 2	LSWT, LSWT²	4.0	3.4	-5.2	3.4
Lake Surprise	LGM	Spring through summer	This study: 3	LSWT, LSWT², φ	3.9	3.0	-5.3	3.0
Lake Surprise	LGM	Spring through summer	This study: 4	LSWT, LSWT², φ, z	3.8	3.0	-5.4	3.0
Lake Surprise	LGM	Spring through summer	Hren and Sheldon (2012)	LSWT, LSWT²	5.1	3.8	-4.1	3.8
Kunlun Pass	Pleistocene	Summer	This study: 1	LSWT	-8.7	5.8	-15.2	5.8
Kunlun Pass	Pleistocene	Summer	This study: 2	LSWT, LSWT²	-4.1	5.5	-10.6	5.5
Kunlun Pass	Pleistocene	Summer	This study: 3	LSWT, LSWT², φ	-4.0	4.4	-10.5	4.4
Kunlun Pass	Pleistocene	Summer	Hren and Sheldon (2012)	LSWT, LSWT²	-6.4	6.1	-12.9	6.1
Kunlun Pass	iNHG	Summer	This study: 1	LSWT	-10.2	5.7	-16.7	5.7
Kunlun Pass	iNHG	Summer	This study: 2	LSWT, LSWT²	-5.1	5.4	-11.6	5.4
Kunlun Pass	iNHG	Summer	This study: 3	LSWT, LSWT², φ	-4.9	4.3	-11.4	4.3
Kunlun Pass	iNHG	Summer	Hren and Sheldon (2012)	LSWT, LSWT²	-7.7	6.0	-14.2	6.0
Kunlun Pass	mPWP	Summer	This study: 1	LSWT	-2.5	6.3	-9.0	6.3
Kunlun Pass	mPWP	Summer	This study: 2	LSWT, LSWT²	-1.9	6.0	-8.4	6.0
Kunlun Pass	mPWP	Summer	This study: 3	LSWT, LSWT², φ	-1.0	5.0	-7.5	5.0
Kunlun Pass	mPWP	Summer	Hren and Sheldon (2012)	LSWT, LSWT²	-0.5	6.5	-7.0	6.5
Kunlun Pass	Pliocene	Summer	This study: 1	LSWT	0.2	6.4	-6.3	6.4
Kunlun Pass	Pliocene	Summer	This study: 2	LSWT, LSWT²	1.1	6.1	-5.4	6.1
Kunlun Pass	Pliocene	Summer	This study: 3	LSWT, LSWT², φ	1.6	5.1	-5.0	5.1
Kunlun Pass	Pliocene	Summer	Hren and Sheldon (2012)	LSWT, LSWT²	1.7	6.7	-4.8	6.7

Table 2: Reconstructed mean annual air temperature predictions and anomalies utilizing transfer functions from this study (transfer functions 1, 2, 3, and 4) and proxy-derived lake surface water temperature (LSWT) reconstructions from Lake Surprise (northwestern Great Basin; Santi et al., 2020) and Kunlun Basin (Tibetan Plateau; Cheng et al., 2022). Estimates derived using the transfer functions developed in Hren and Sheldon (2012) are also shown for comparison.

Acknowledgements

We thank Karen McKinnon for feedback on an early version of this work and the Tripati Lab group for their support. We also thank Stephen Cropper for their feedback over the course of this project. This work was supported by NSF CAREER award EAR-1352212, NSF ICER-1936715, NSF RISE-2228198, and Heising-Simons Foundation award 2021-3137. Alexa Terrazas, Alexandrea Arnold, Robert Ulrich, and Aradhna Tripati were supported by the Center for Diverse Leadership in Science, which is funded by the Waverley Street Foundation, Silicon Valley Community Foundation, Packard Foundation, Sloan Foundation. They were also supported by a Royal Society Wolfson Visiting Research Fellowship. Alexa Terrazas and Robert Ulrich were supported by the NSF Graduate Research Fellowship Program; Alexandrea Arnold was supported by a UC Dissertation Year Fellowship and a Cota-Robles Fellowship.

Author Contribution

A.T. and A.J.A conceived of the study. A.T and N.H. conducted the analysis. A.J.A, and A.T., downloaded the data. A.T and N.H. consulted R.N.U for the machine learning component. A.E.T. assisted A.Tripati. in the drafting of the manuscript. All co-authors read and provided comments and substantive revisions to the manuscript.

Data Availability Statement

The code will be made available on https://github.com/Tripati-Lab upon publication. Data is accessible in the Supporting documents section.

Conflict of Interest

None declared.