S1 Introduction

This supplement contains

- Additional information on the ensemble of transient simulations of the last deglaciation
- Additional information on the proxy record pre-processing
- Details on the approximation of the Integrated Quadratic Distance using Monte Carlo samples and the sensitivity of results on the number of Monte Carlo samples
- Exemplary plots of time series decompositions

S2 Additional information on simulation ensemble

S2.1 MPI-ESM simulations

The MPI-ESM simulations use the coarse resolution version of MPI-ESM v.1.2 (MPI-ESM-CR). The atmospheric component runs at T31 horizontal resolution ($\sim 3.75°$) and 31 vertical levels. The ocean component employs an unstructured grid with a nominal resolution of $3°$ (Kapsch et al., 2022). Orbital parameters follow Berger and Loutre (1991) and greenhouse gas concentrations follow Köhler et al. (2017). Ice sheet topographies, surface topographies, glacier masks, land-sea masks, and bathymetries are updated based on either the GLAC-1D or the ICE-6G histories as described in the main manuscript and Kapsch et al. (2022). Ice sheet topography, glacier mask, land-sea mask, and bathymetry fields are interpolated to a 10-year resolution and updated every 10 simulation years (Kapsch et al., 2022). Accordingly, river pathways are updated every 10 simulation years based on the surface fields. Meltwater fluxes are computed as the temporal derivative of ice sheet thickness in the respective grid boxes (Kapsch et al., 2022). In the simulations with local meltwater input (see Table 1 in the main manuscript), the locations and amounts of meltwater are determined with a dynamically adapting river runoff scheme (Riddick et al., 2018). In MPI_Ice6G_P2_glob, meltwater is distributed equally across all land and ocean grid cells. In MPI_Ice6G_P2_noMWF, meltwater is removed from the system. Vegetation cover changes dynamically as computed by the dynamic global vegetation model JSBACH which is incorporated in MPI-ESM (Kapsch et al., 2022).

S2.2 CCSM3 simulations

The TraCE simulation design is described in detail in He (2011). The simulations are run with CCSM3 using a T31 horizontal resolution ($\sim 3.75°$) with 26 vertical levels for the atmosphere and a variable horizontal resolution with 25 vertical levels for the ocean ($3.6°$ in longitudinal direction, $\sim 0.9°$ in latitudinal direction near the equator but coarser at higher latitudes). Orbital parameters follow Berger (1978). Greenhouse gas concentrations are changed according to Joos and Spahni (2008). Ice sheet masks and topographies are updated approximately every 500 yrs following the ICE-5G history (see He, 2011, for details). Land-sea masks and bathymetries follow the ICE-5G reconstruction but are only adapted at four time steps at 13.1 ka (Barents Sea opens), 12.9 ka (Bering Strait opens), 7.6 ka (Hudson Bay opens), and 6.2 ka (Indonesian Throughflow begins). Location and amounts of meltwater fluxes were adapted manually as described in He (2011). Vegetation cover changes dynamically in the integrated dynamic global vegetation model CLM-DGVM (He, 2011). In TraCE-GHG and TraCE-ORB only greenhouse gas concentrations and orbital parameters are changed respectively, whereas all other boundary conditions are kept at the LGM state (22 ka) of the TraCE-ALL simulation.

S2.3 FAMOUS simulation

FAMOUS is a coarse resolution and simplified version of the HadCM3 climate model (Smith, 2012). The atmosphere module has a horizontal grid spacing of $5° \times 7.5°$ with 11 vertical levels and the ocean module has a horizontal grid spacing of $2.5° \times 3.75°$ with 20 vertical levels (Smith, 2012). In the assessed simulation, orbital parameters are changed transiently following Berger (1978) and greenhouse gas concentrations follow Lüthi et al. (2008). Northern Hemisphere (north of $40°$N) ice sheet topographies and glacier masks are updated according to the ICE-5G ice sheet history, whereas the Antarctic ice sheet topography and glacier mask, the land-sea mask, and the ocean bathymetry are fixed at pre-industrial values (Smith and Gregory, 2012). Due to the unchanged land-sea mask, ice sheet topographies are only prescribed for pre-industrial land grid cells and not over pre-industrial ocean grid cells. Meltwater fluxes are removed from the system. The vegetation cover is fixed at pre-industrial values. All boundary condition changes are applied with a 10x acceleration factor (Smith and Gregory, 2012).

S3 Details of proxy record processing

The PalMod 130k marine paleoclimate data synthesis v1.1.1 contains (near-)surface temperature reconstructions from original publications (Jonkers et al., 2020, 2023). While age models were recomputed and harmonized as described in Jonkers et al. (2020), the SST calibrations proposed by the original authors are used. We build on the published reconstructions, provided

in a harmonized and metadata rich format in the database. The database combines temperature reconstructions from multiple sensors. In particular, the selected records contain (near-)surface temperature reconstructions from Mg/Ca, U^k_{37}, planktonic foraminifera assemblages, TEX_{86}, and diatom assemblages. For some records, multiple calibrations of the same sensor measurements are archived in the database. To avoid biases from including sensor measurements multiple times, a preprocessing of the data provided in the database is required.

This preprocessing consists of five steps. First, we collect all available time series of (near-)surface temperature reconstructions in the database according to the metadata parameter 'surface.temp'. This results in 252 time series from 132 unique marine sediment cores.

Second, we remove records without BACON ages at any depth with available (near-)surface temperature reconstructions. This case can occur because BACON chronologies are only available for the sample depths between the first and last dating point. This step results in 247 time series from 132 unique marine sediment cores.

Third, we harmonize the provided information on recording or calibration season of the sensors/reconstructions. For some reconstructions, in particular many of the planktonic foraminifera assemblage reconstructions, a recording or calibration season is provided in the original data publication. This information is stored either in the ParameterOriginal or the RecordingSeason parameter of the database. We categorize the reconstructed time series into 'annual', 'warm season', 'cold season', or 'unknown' based on the available metadata. We only use the ParameterOriginal if no RecordingSeason is provided. 'unknown' is assigned whenever neither ParameterOriginal nor RecordingSeason contain information on the recording or calibration season.

Fourth, we average warm and cold season temperature reconstructions from the same core and sensor to obtain pseudo-annual temperature reconstructions. In all cases where warm and cold season temperature reconstructions are available, the same number of time series is available for each season. Therefore, we can average them without the need to first create averages for each season. This step results in 205 remaining time series from 132 sediment cores.

Fifth, we average and select records such that we do not include multiple reconstructed time series from the same sensor and core (Mg/Ca temperatures derived from different species are treated as different sensors). We average pseudo-annual, annual, and unknown seasonality reconstructions from the same core and sensor if all sample depths coincide. If the sample depths of time series from the same core and sensor do not coincide, we manually select the time series that provides a better coverage of the period 22-6 ka. Here, the length of the covered interval within that period and the temporal resolution are the determining criteria. In this process, we also remove one time series from core 'SU81_18' which has an unknown sensor in the database. Resolving these special cases results in 186 remaining time series from 132 sediment cores.

Using the record quality criteria described in the main manuscript (Sect. 2.2), we end up with 74 time series from 50 sediment cores which are used in the analyses of this manuscript.

Note that the correlations between time series calibrated from the same measurements are mostly high due to an often similar temporal pattern of orbital- and millennial-scale temperature variations. Therefore, the influence of the averaging on the pattern of orbital- and millennial-scale variations tends to be small. Calibrations for different seasons often vary in the magnitudes of variations. The construction of pseudo-annual records averages between these different magnitudes (since the temporal patterns for the different seasons are often highly correlated). Reconstructions for the same season tend to have similar magnitudes of variations such that the effect of averaging multiple calibrations for the same season is mostly small.

The preprocessing procedure is reproducible using the code provided in support of this manuscript (see section 'Code and data availability' in the main manuscript).

S4 Computational details

S4.1 Approximation of the Integrated Quadratic Distance using Monte Carlo samples

To compute the Integrated Quadratic Distance (IQD, Sect. 3.1.3 of the main manuscript), we need to approximate

$$\mathrm{IQD}(\mathbb{P},\mathbb{Q}) = \frac{1}{M}\mathbb{E}_{\mathbb{P},\mathbb{Q}}|X-Y| - \frac{1}{2M}\left(\mathbb{E}_{\mathbb{P}}|X-X'| + \mathbb{E}_{\mathbb{Q}}|Y-Y'|\right), \tag{S1}$$

with the N Monte Carlo (MC) samples of reconstructed and forward-modeled proxy time series. Remember that $\mathbb{P}$ is the probability distribution of forward-modeled proxy time series, $\mathbb{Q}$ is the probability distribution of the reconstructions, M is the

dimension of $\mathbb{P}$ and $\mathbb{Q}$, and $\mathbb{E}$ denotes expected values. Further, X and X' are independent random variables distributed according to $\mathbb{P}$, and Y and Y' are independent random variables distributed according to $\mathbb{Q}$. A standard Monte Carlo approximation of Equ. (S1) is

$$\mathrm{IQD}(\mathbb{P},\mathbb{Q}) \approx \frac{1}{MN}\sum_{n=1}^{N}|X_n - Y_n| - \frac{1}{2M}\frac{2}{N(N-1)}\left(\sum_{n=2}^{N}\sum_{m<n}|X_n - X_m| + \sum_{n=2}^{N}\sum_{m<n}|Y_n - Y_m|\right), \tag{S2}$$

where $|\cdot|$ denotes the Euclidean distance between one- or multi-dimensional vectors and X_n, X_m, Y_n and Y_m are the respective MC samples. Equ. (S2) needs $O(N^2)$ operations which becomes computationally expensive for high dimensional vectors such as the pattern time series. Therefore, we employ a less accurate approximation of the last two terms in Equ. (S1), which only needs $O(N)$ operations:

$$\mathrm{IQD}(\mathbb{P},\mathbb{Q}) \approx \frac{1}{MN}\sum_{n=1}^{N}|X_n - Y_n| - \frac{1}{2MN}\left(\sum_{n=1}^{N-1}|X_n - X_{n+1}| + |X_N - X_1| + \sum_{n=1}^{N-1}|Y_n - Y_{n+1}| + |Y_N - Y_1|\right). \tag{S3}$$

We find that using the less accurate Equ. (S3) with larger N is in total more accurate than using Equ. (S2) with lower N (not shown). Therefore, we use Equ. (S3) to approximate Equ. (S1).

S4.2 Sensitivity to number of Monte Carlo samples

To find out how many Monte Carlo (MC) samples are needed to obtain stable IQD approximations, we perform pseudo-proxy experiments (PPEs).

We use four different simulations as reference simulations and employ the proxy system model (PSM) with an AR1 noise using SNR=1.6 and a decorrelation length of 1289 yrs. The four reference simulations are MPI_Glac1D_P3, MPI_Ice6G_P2_glob, TraCE-ALL, and FAMOUS.

To compute the IQD, we use 5, 10, 25, 50, 75, 100, 250, 500, 750, and 1000 MC samples. To isolate the effect from including additional MC samples, we always use all MC samples from the next smaller sample size and then include additional samples.

In Fig. S1, we plot the mean change in absolute IQD per 10 additional MC samples. For larger sample sizes, absolute IQD differences per 10 additional MC samples become smaller. In particular, IQD differences per additional MC samples become negligible when increasing the MC sample size beyond $N = 100$. This finding holds for all four components of the deglacial temperature evolution and all three spatial averaging scales (global, zonal, and local). We conclude that IQD estimates are robust for MC sample sizes of $N \geq 100$. Therefore, we employ $N = 100$ in the computationally expensive large sets of PPEs (used in Sect. 4.1.2 and 4.1.3 of the main manuscript), but use $N = 1000$ in the example PPE (Sect. 4.1.1) and the real-world application (Sect. 4.2).

S5 Additional time series decompositions

We plot additional examples of time series decompositions into orbital, millennial, and sub-millennial contributions (compare with Fig. 2 of the main manuscript). For each of the selected 74 proxy records from the PalMod 130k marine paleoclimate data synthesis v1.1.1 (Jonkers et al., 2020, 2023) (Sect. 2.2 of the main manuscript), we randomly select an age-depth history and a realization of the PSM (SNR = 1.6, AR1 noise with decorrelation length 1289 yrs) with MPI_Glac1D_P3 as reference simulation. We derive the orbital contributions for three smoothing periods, 4 kyr, 6 kyr, and 8 kyr (this leads to three versions of the orbital variations). Millennial and sub-millennial variations are separated using a smoothing period of 1 kyr. Therefore the sum of orbital and millennial variations is the same in all three decompositions. Fig. S2 - Fig. S9 show this single realization of the reconstructed and the forward-modeled proxy time series (black), the orbital variations (blue), and the sum of orbital and millennial variations (red) for each of the 74 proxy records.

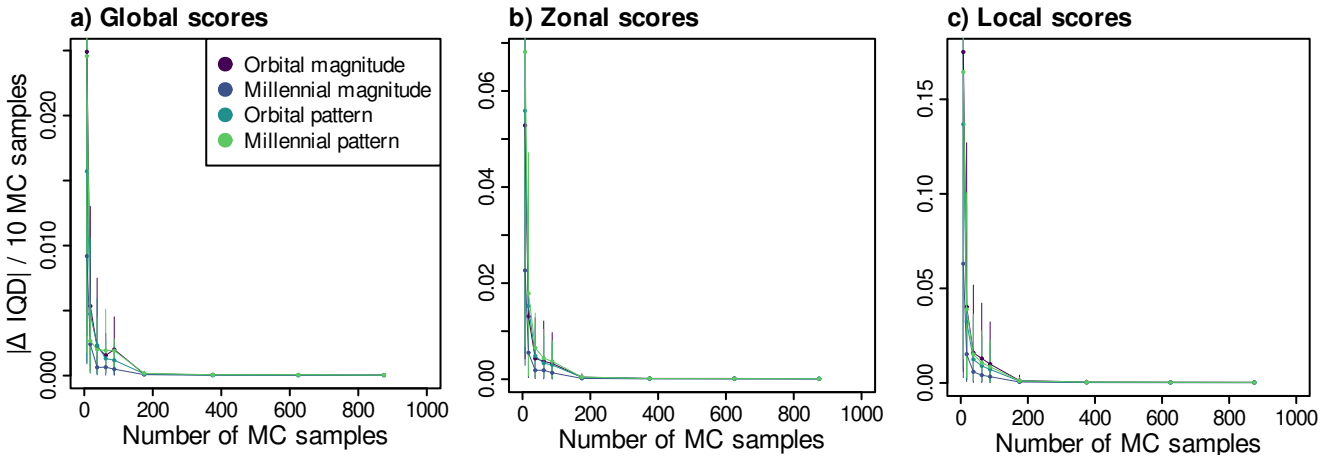

Figure S1. Changes in the absolute IQDs per 10 additional MC samples for (a) globally averaged IQDs, (b) zonally averaged IQDs, and (c) local IQDs. Values are positive by definition. Connected dots correspond to means over all PPEs with the same number of MC samples, whereas vertical lines denote 90% confidence intervals across PPEs with the same number of MC samples (given by the 5th and 95th percentiles of the respective distributions). As each dot corresponds to the difference in IQD between subsequently tested sample sizes, dots are plotted at the mean of the two sample sizes.

References

Berger, A.: Long-Term Variations of Daily Insolation and Quaternary Climatic Changes, J. Atmos. Sci., 35, 2362–2367, https://doi.org/10.1175/1520-0469(1978)035<2362:LTVODI>2.0.CO;2, 1978.

Berger, A. and Loutre, M.: Insolation values for the climate of the last 10 million years, Quaternary Science Reviews, 10, 297–317, https://doi.org/10.1016/0277-3791(91)90033-Q, 1991.

He, F.: Simulating Transient Climate Evolution of the Last Deglaciation with CCSM3, Dissertation, University of Wisconsin-Madison, 2011.

Jonkers, L., Cartapanis, O., Langner, M., McKay, N., Mulitza, S., Strack, A., and Kucera, M.: Integrating palaeoclimate time series with rich metadata for uncertainty modelling: strategy and documentation of the PalMod 130k marine palaeoclimate data synthesis, Earth Syst. Sci. Data, 12, 1053–1081, https://doi.org/10.5194/essd-12-1053-2020, 2020.

Jonkers, L., Cartapanis, O., Langner, M., McKay, N., Mulitza, S., Strack, A., and Kucera, M.: PalMod 130k marine palaeoclimate data synthesis version 1.1.1, https://doi.org/10.5281/ZENODO.7785766, 2023.

Joos, F. and Spahni, R.: Rates of change in natural and anthropogenic radiative forcing over the past 20,000 years, Proc. Natl. Acad. Sci. U.S.A., 105, 1425–1430, https://doi.org/10.1073/pnas.0707386105, 2008.

Kapsch, M., Mikolajewicz, U., Ziemen, F., and Schannwell, C.: Ocean Response in Transient Simulations of the Last Deglaciation Dominated by Underlying Ice-Sheet Reconstruction and Method of Meltwater Distribution, Geophysical Research Letters, 49, https://doi.org/10.1029/2021GL096767, 2022.

Köhler, P., Nehrbass-Ahles, C., Schmitt, J., Stocker, T. F., and Fischer, H.: A 156 kyr smoothed history of the atmospheric greenhouse gases CO_2, CH_4, and N_2O and their radiative forcing, Earth Syst. Sci. Data, 9, 363–387, https://doi.org/10.5194/essd-9-363-2017, 2017.

Lüthi, D., Le Floch, M., Bereiter, B., Blunier, T., Barnola, J.-M., Siegenthaler, U., Raynaud, D., Jouzel, J., Fischer, H., Kawamura, K., and Stocker, T. F.: High-resolution carbon dioxide concentration record 650,000–800,000 years before present, Nature, 453, 379–382, https://doi.org/10.1038/nature06949, 2008.

Riddick, T., Brovkin, V., Hagemann, S., and Mikolajewicz, U.: Dynamic hydrological discharge modelling for coupled climate model simulations of the last glacial cycle: the MPI-DynamicHD model version 3.0, Geosci. Model Dev., 11, 4291–4316, https://doi.org/10.5194/gmd-11-4291-2018, 2018.

Smith, R. S.: The FAMOUS climate model (versions XFXWB and XFHCC): description update to version XDBUA, Geosci. Model Dev., 5, 269–276, https://doi.org/10.5194/gmd-5-269-2012, 2012.

Smith, R. S. and Gregory, J.: The last glacial cycle: transient simulations with an AOGCM, Clim Dyn, 38, 1545–1559, https://doi.org/10.1007/s00382-011-1283-y, 2012.

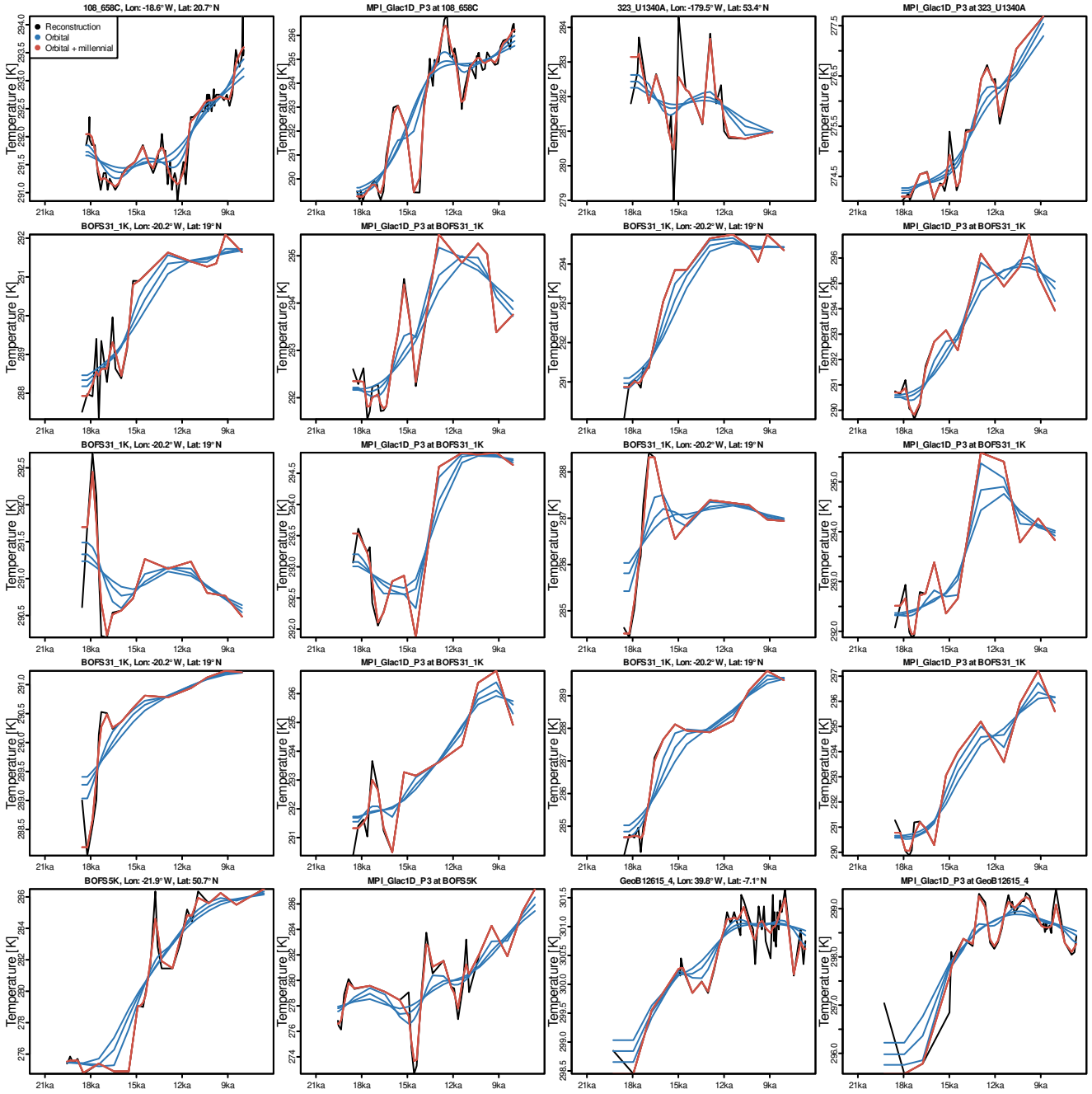

Figure S2. Single realization of the reconstructed and the forward-modeled proxy time series (black), the orbital variations (blue), and the sum of orbital and millennial variations (red) for each of the 74 proxy records (see Sect. S5 for details). Orbital and millennial variations are separated using three different smoothing periods, 4 kyr, 6 kyr, and 8 kyr, leading to three versions of the orbital variations with different degrees of smoothing.

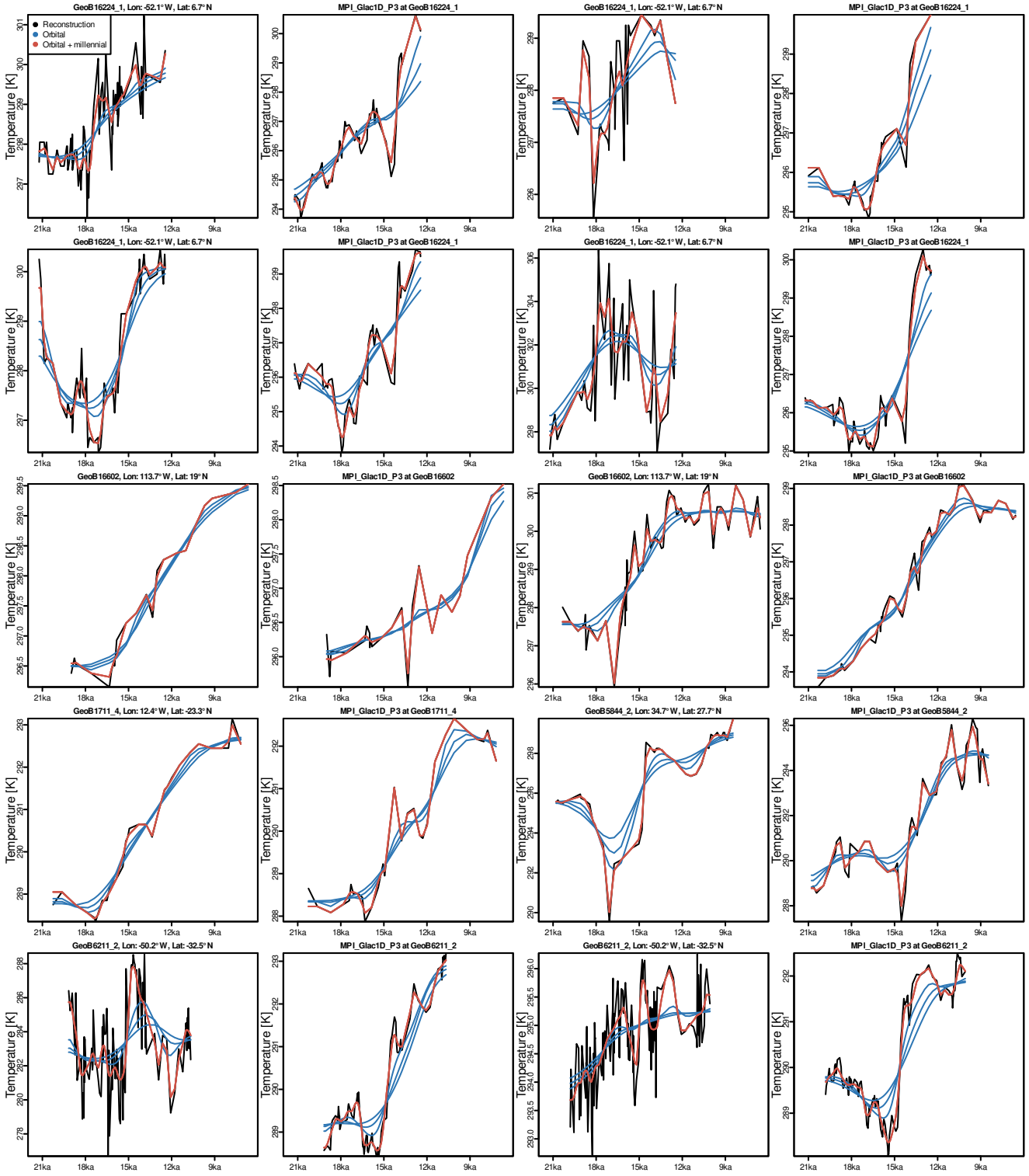

Figure S3. As Fig. S2 but for other proxy records.

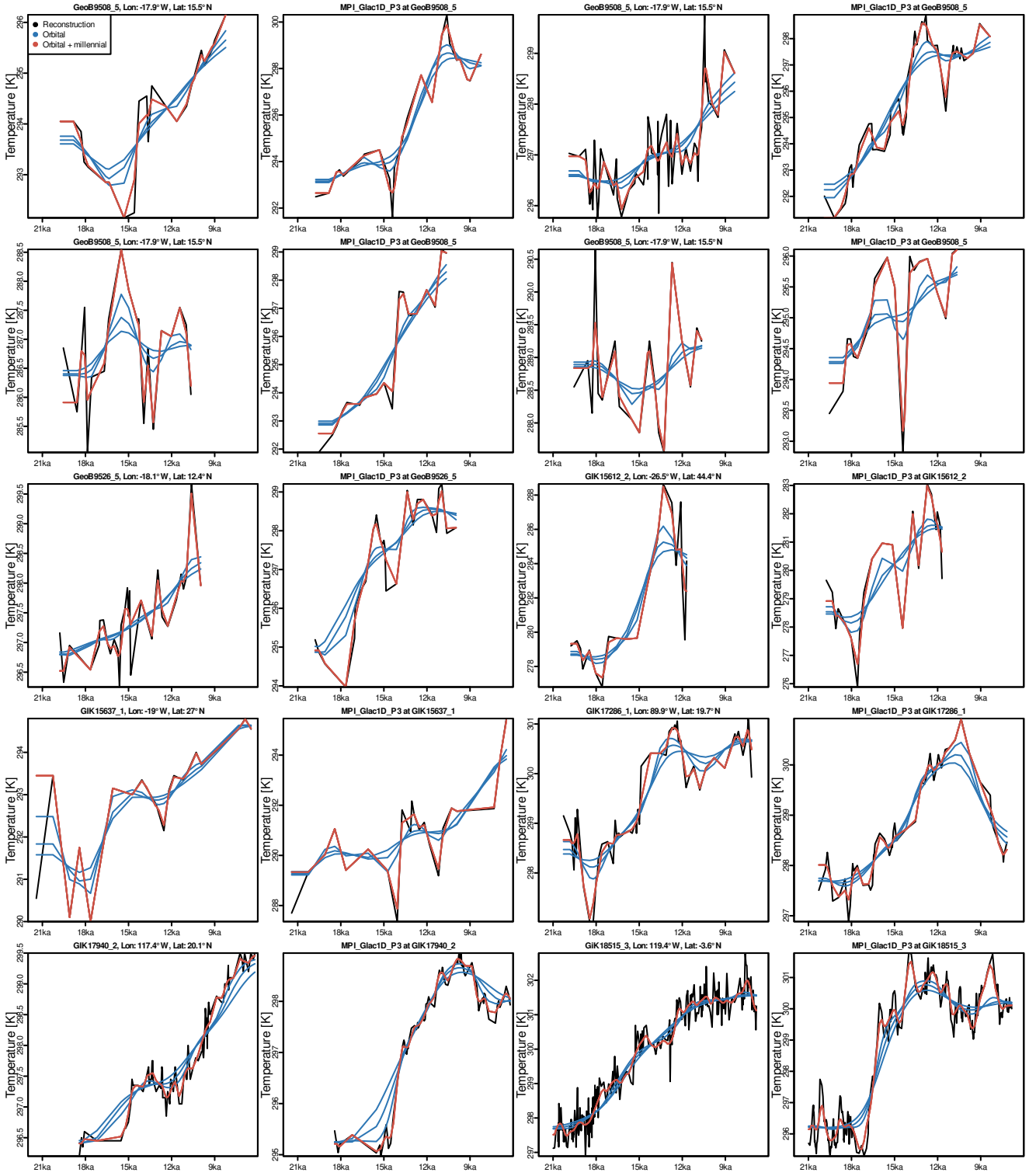

Figure S4. As Fig. S2 but for other proxy records.

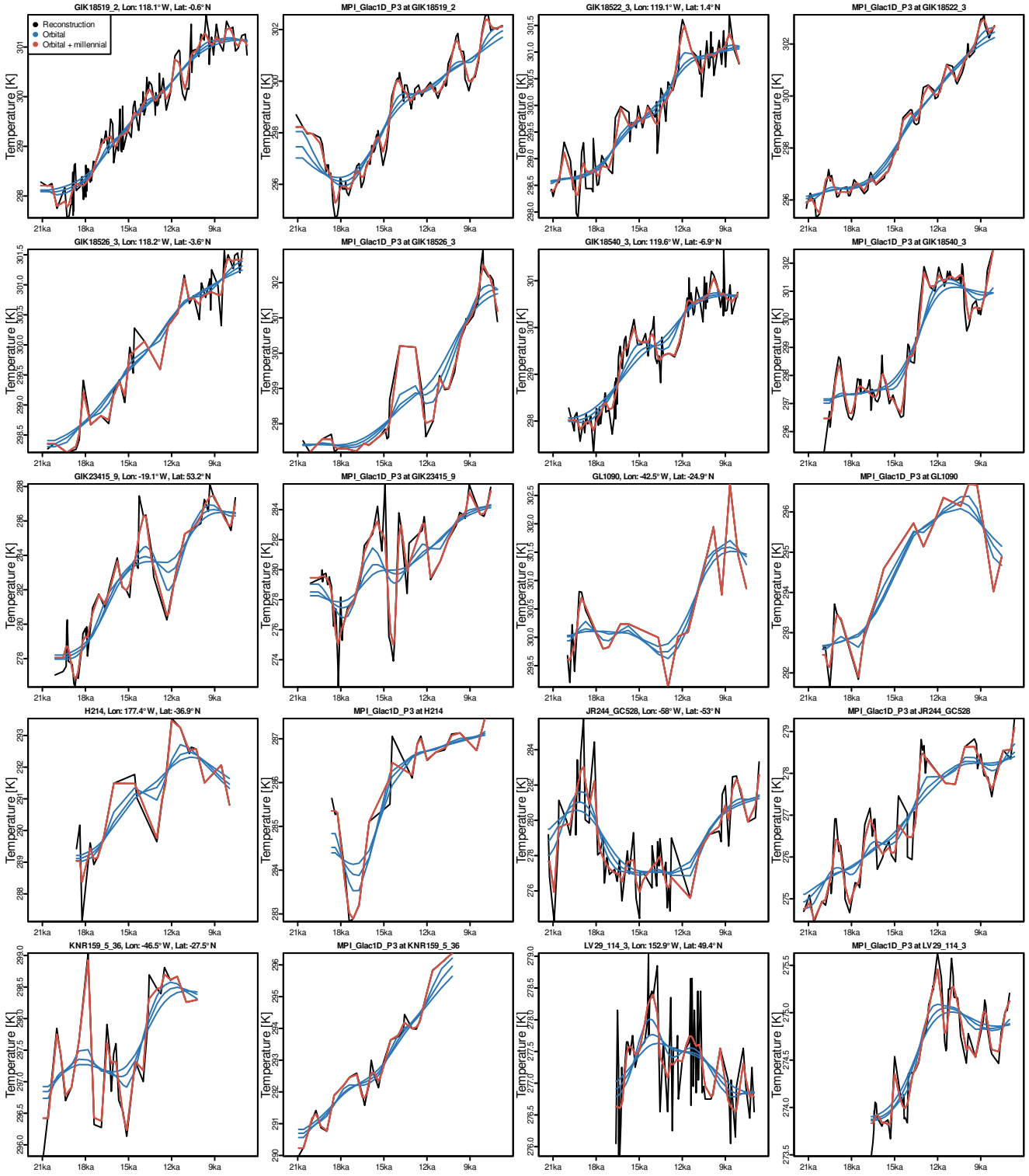

Figure S5. As Fig. S2 but for other proxy records.

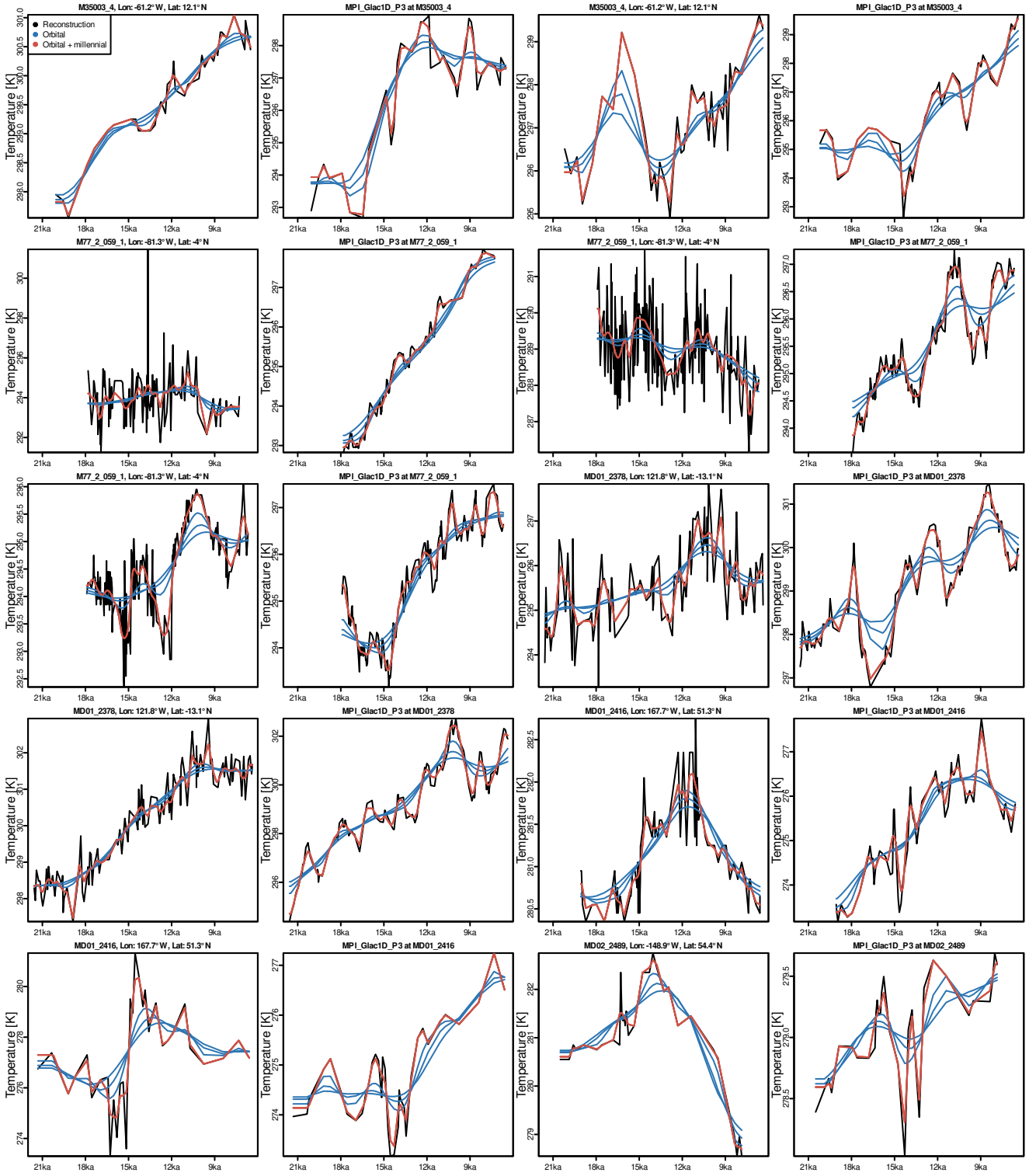

Figure S6. As Fig. S2 but for other proxy records.

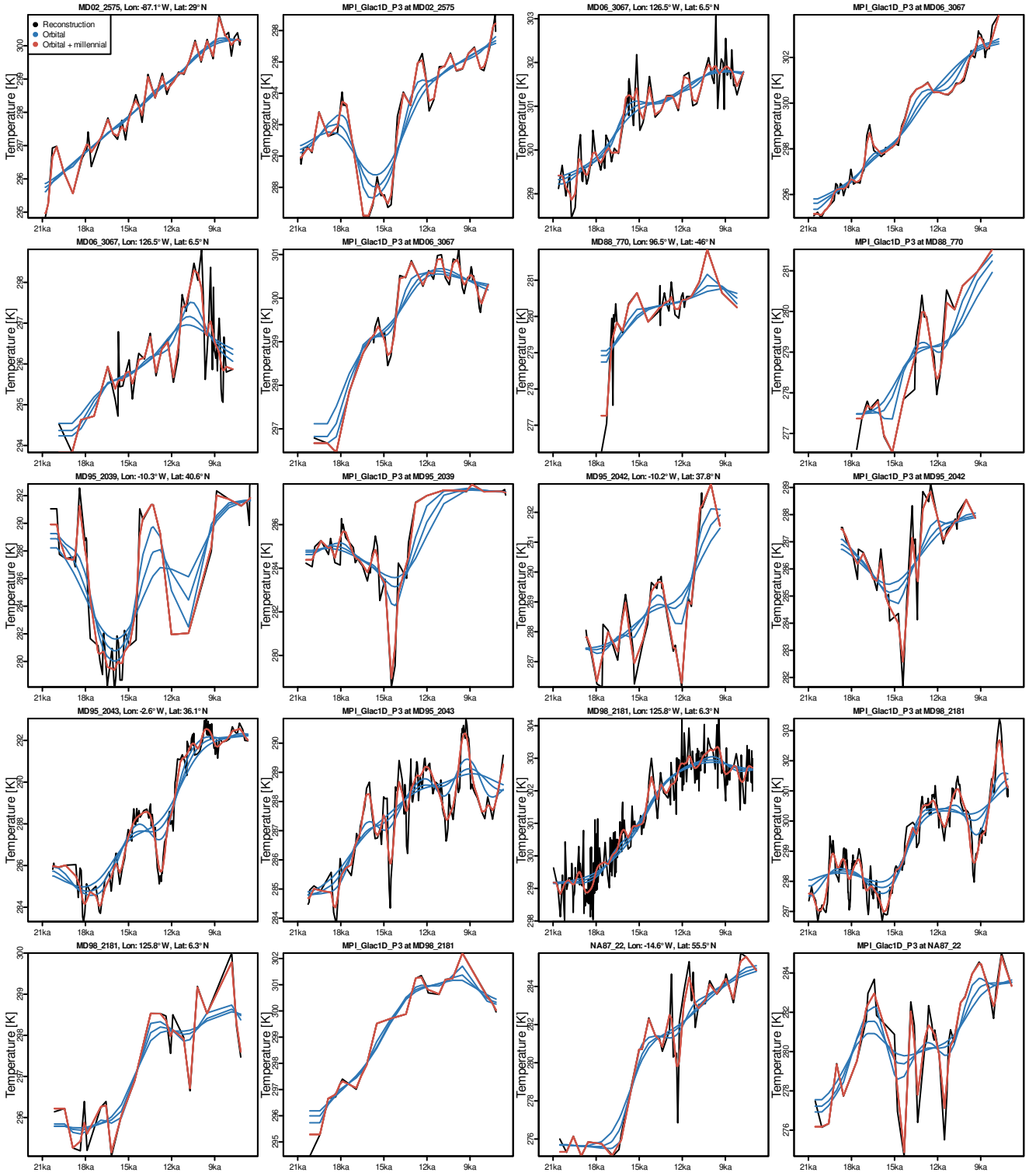

Figure S7. As Fig. S2 but for other proxy records.

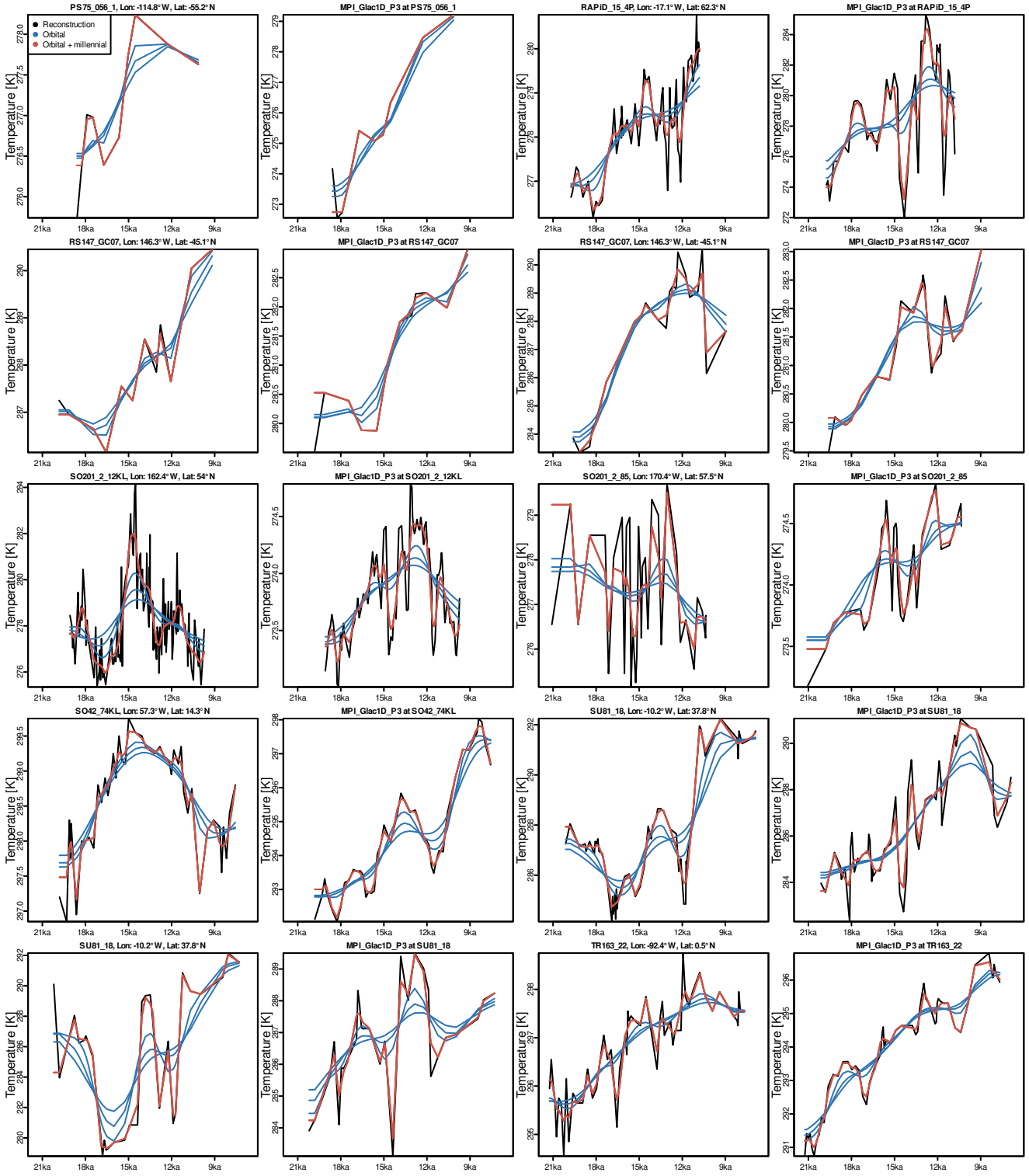

Figure S8. As Fig. S2 but for other proxy records.

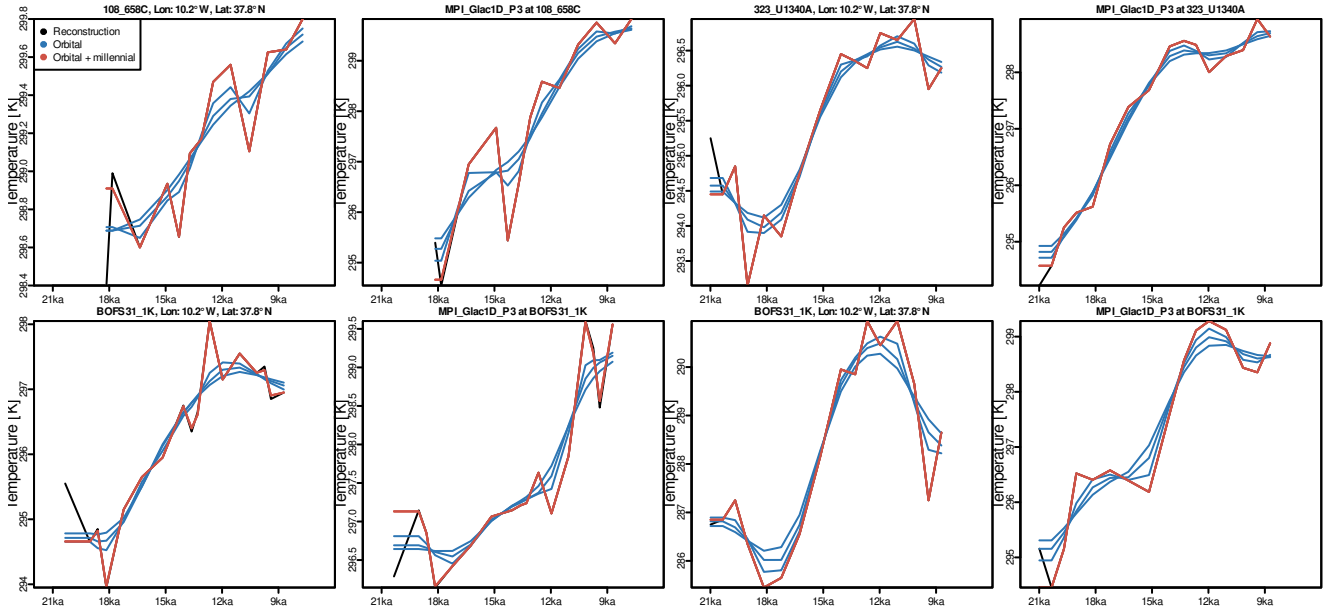

Figure S9. As Fig. S2 but for other proxy records.