Supplement of Geochronology, 6, 503–519, 2024
https://doi.org/10.5194/gchron-6-503-2024-supplement

Supplement of

Towards the construction of regional marine radiocarbon calibration curves: an unsupervised machine learning approach

Ana-Cristina Mârza et al.

Correspondence to: Luke C. Skinner (lcs32@cam.ac.uk)

SUPPLEMENTARY MATERIAL

Table S1 summarises the most important software packages used in our analysis. The scripts and complete python environment specifications are hosted at: https://gitlab.com/earth15/ocean_data_clusters.

Table S1: Principal python packages used.

Package name	**Version**	**Usage**	**Reference**
python	3.9.14	Main scripting language	("The Python Language Reference — Python 3.9.14 Documentation")
numpy	1.22.4	Data manipulation	Harris et al. (2020)
matplotlib	3.5.3	Data visualisation	Hunter (2007)
nctoolkit	0.8.4	NetCDF file manipulation, spatiotemporal statistics	https://nctoolkit.readthedocs.io/en/latest/info.html
sktime	0.14.0	Time series K-Medoids	Löning et al. (2019)
tslearn	0.5.2	Time series K-Means, time series normalisation	Tavenard et al. (2020)
kneed	0.8.2	Locating the elbow point in a plot	https://kneed.readthedocs.io/en/stable/index.html
scipy	1.9.3	Hierarchical clustering	Virtanen et al. (2020)
xesmf	0.6.3	Regridding (nearest_s2d algorithm)	Zhuang et al. (2022)

Figure S1. *Clustering results using K-medoids on un-normalised data from the CM2Mc interglacial run. As for Figure 4 of the main text, but for K increased to K=9.*

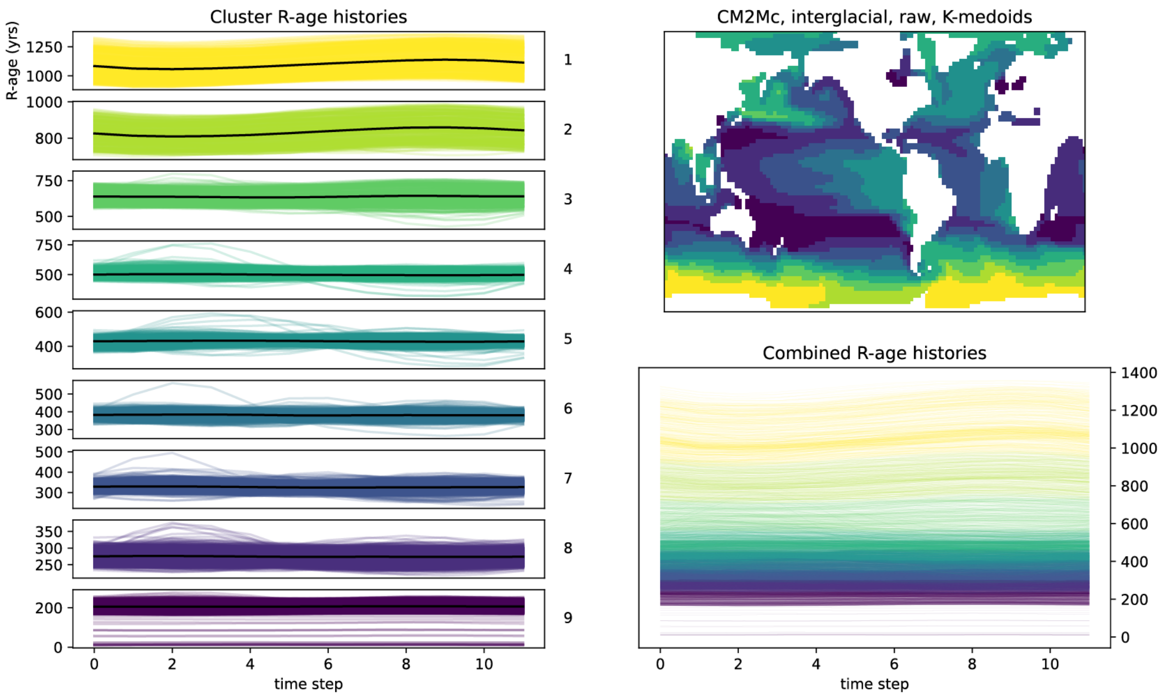


Figure S2. *Clustering results using K-Medoids on UVic (U-Tr run) normalised data. As for Figure 8 of the main text, but for K increased to K=8.*

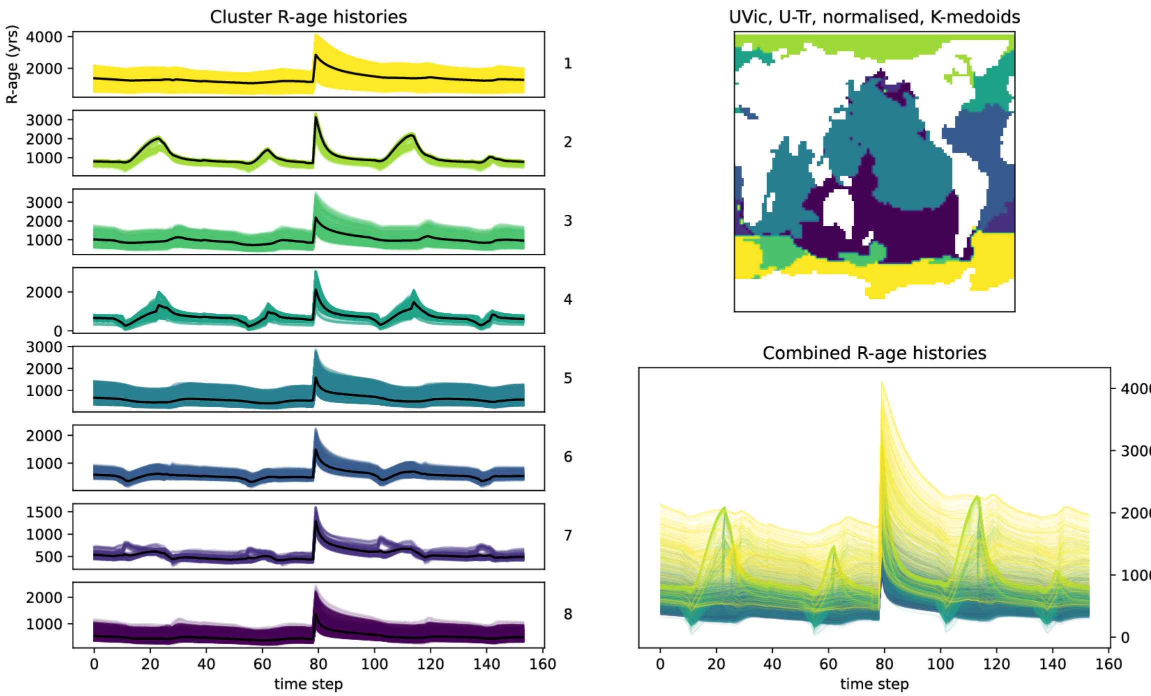


Figure S3. *Comparison of clustering results between the two UVic runs (left: U-Tr; right: U-TrS), using K-medoids on normalised data. The colour-coded Venn diagrams quantify overlap between the two maps. Note the strong agreement between the two resulting cluster geographies*

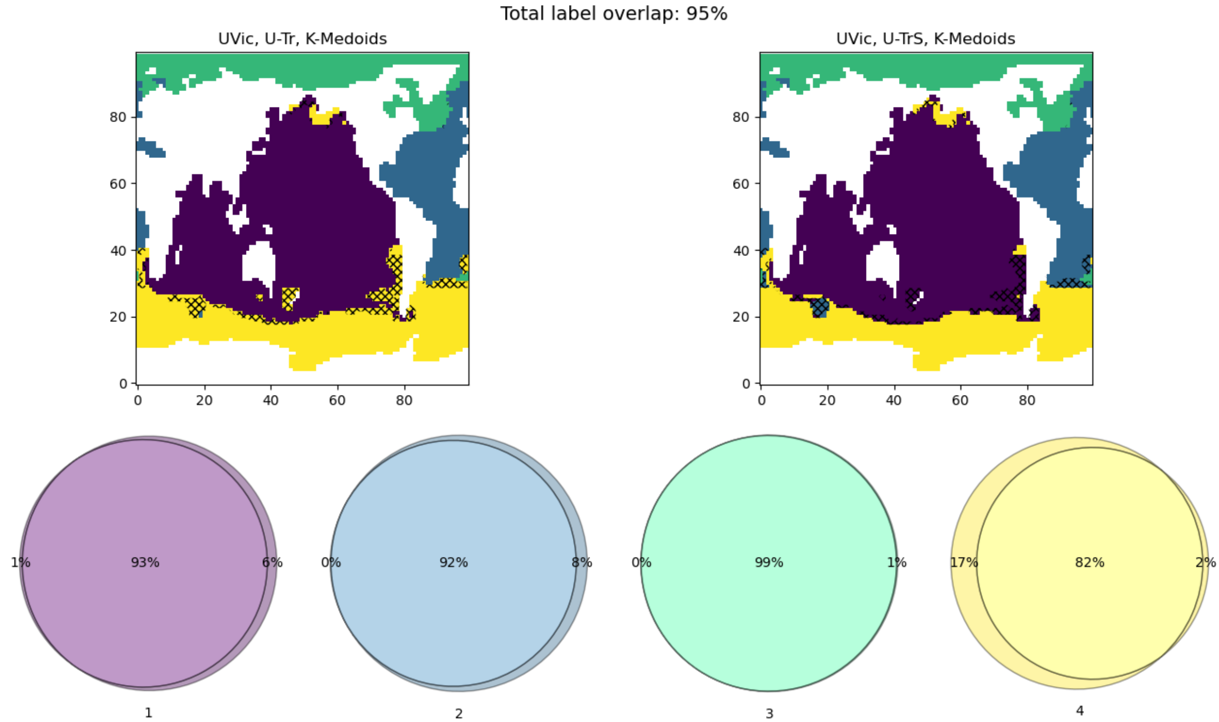


Figure S4. *The variance (a.) and mean (b.) of the distribution of R-age offsets in CM2Mc glacial data, when computed relative to three different references: the running mean of the global surface ocean, analogous to e.g. Marine20 (column 1); the medoid of the shape-based cluster to which the subcluster belongs (column 2); and the subcluster's own medoid (column 3). Crucially, and in contrast to Figure 13 of the main text, the geographic definitions of the clusters and their medoids are derived from a different model run (UVic U-Tr). UVic cluster labels were re-gridded to match the resolution of the CM2Mc data.*

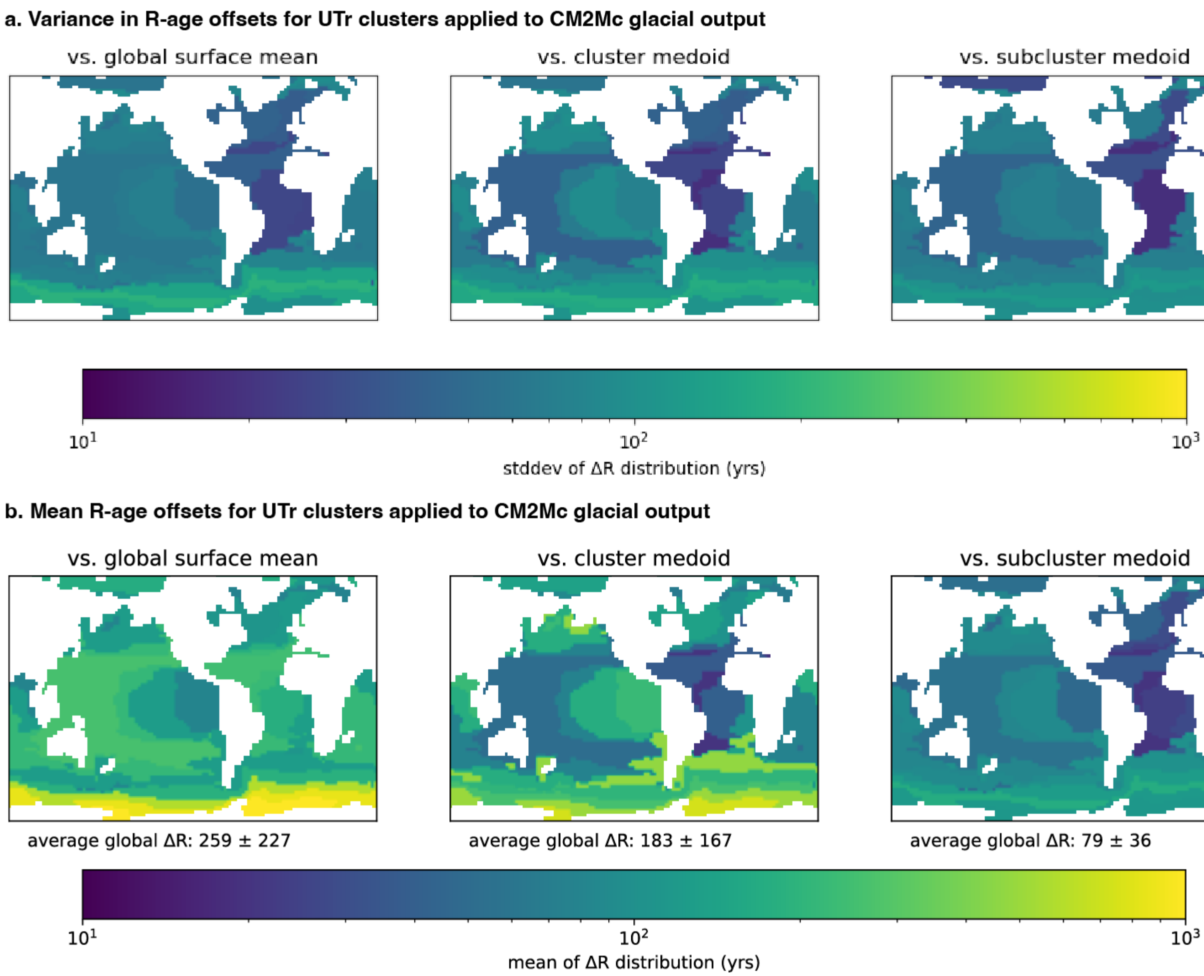
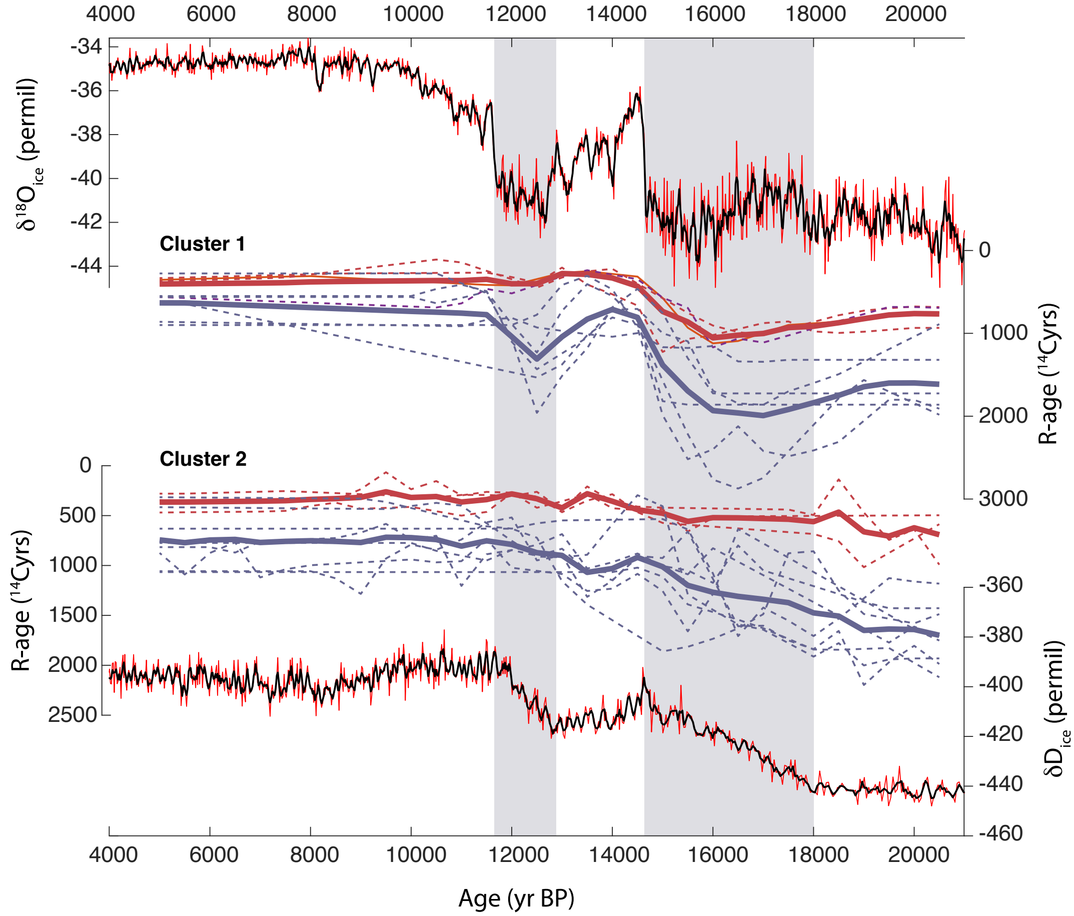


Figure S5. *Tentative hierarchical clustering of R-age reconstructions spanning the last deglaciation (Skinner et al., 2023). Plot a: Greenland ice-core temperature proxy, NGRIP $\delta^{18}O_{ice}$ (heavy black line is 3-point running mean). Plot b: amplitude-based sub-clusters (K = 2, red and blue lines) belonging to shape-based Cluster 1; heavy lines indicate the sub-cluster averages and dashed lines indicate the clustered time-series data. Plot c, as for plot b, but for sub-clusters within the shape-based Cluster 2. Plot d: Antarctic ice-core temperature proxy, EDC δD_{ice} (heavy black line is 3-point running mean). Vertical bars indicate approximate timing of North Atlantic stadials corresponding to the Younger Dryas and Heinrich Stadial 1.*



References

Harris, C. R., Millman, K. J., Walt, S. J. v. d., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N. J., Kern, R., Picus, M., Hoyer, S., Kerkwijk, M. H. v., Brett, M., Haldane, A., Río, J. F. d., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., and Oliphant, T. E.: Array programming with NumPy, Nature, 585, 357–362, 10.1038/s41586-020-2649-2, 2020.

Hunter, J. D.: Matplotlib: A 2D graphics environment, Computing in Science & Engineering, 9, 90–95, 10.1109/MCSE.2007.55, 2007.

Löning, M., Bagnall, A., Ganesh, S., Kazakov, V., Lines, J., and Király, F. J.: sktime: A Unified Interface for Machine Learning with Time Series, 2019.

Skinner, L., Primeau, F., Jeltsch-Thommes, A., Joos, F., Koehler, P., and Bard, E.: Rejuvenating the ocean: mean ocean radiocarbon, CO2 release, and radiocarbon budget closure across the last deglaciation, Clim. Past, 2023.

Tavenard, R., Faouzi, J., Vandewiele, G., Divo, F., Androz, G., Holtz, C., Payne, M., Yurchak, R., Rußwurm, M., Kolar, K., and Woods, E.: Tslearn, A Machine Learning Toolkit for Time Series Data, Journal of Machine Learning Research, 21, 1-6, 2020.

Virtanen, P., Gommers, R., Oliphant, T. E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., van der Walt, S. J., Brett, M., Wilson, J., Millman, K. J., Mayorov, N., Nelson, A. R. J., Jones, E., Kern, R., Larson, E., Carey, C. J., Polat, İ., Feng, Y., Moore, E. W., VanderPlas, J., Laxalde, D., Perktold, J., Cimrman, R., Henriksen, I., Quintero, E. A., Harris, C. R., Archibald, A. M., Ribeiro, A. H., Pedregosa, F., van Mulbregt, P., and Contributors, S.: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python, Nature Methods, 17, 261–272, 10.1038/s41592-019-0686-2, 2020.

Zhuang, J., dussin, r., Huard, D., Bourgault, P., Banihirwe, A., Raynaud, S., Malevich, B., Schupfner, M., Levang, S., Jüling, A., Almansi, M., Filipe, RichardScottOZ, RondeauG, Rasp, S., Stachelek, J., Bell, R., Smith, T. J., and Li, X.: pangeo-data/xESMF: v0.7.0, Zenodo [code], 2022.