

Supplementary information

European forest cover during the Holocene reconstructed from pollen records

Luke Sweeney^{1,2,*}, Sandy P. Harrison^{1,2}, Marc Vander Linden³

1: Geography and Environmental Science, School of Archaeology, Geography and Environmental

5 Science (SAGES), University of Reading, Whiteknights, Reading, RG6 6AH, UK. Email:
l.sweeney@pgr.reading.ac.uk

2: Leverhulme Centre for Wildfires, Environment and Society, Imperial College London, South
Kensington, London, SW7 2BW, UK. Email: s.p.harrison@reading.ac.uk

3: Institute for the Modelling of Socio-Environmental Transitions, Bournemouth University,
10 Christchurch House, Talbot Campus, Poole, BH 12 5BB, UK. Email:
mvanderlinden@bournemouth.ac.uk

*Correspondence to: Luke Sweeney (l.sweeney@pgr.reading.ac.uk)

Contents

15 This supplement contains the following sections:

- S1: Species categorisation
- S2: Coefficients and model fit based on lake sites only
- S3: Coefficients and model fit without exclusion of sites >1000 m
- S4: Quantile mapping influence on model predictions
- 20 • S5: Spatial structure of model predictions compared to observations
- S6: Modern predictions compared with other reconstructions based on grid cell averages
- S7: Mean reconstructed tree cover
- S8: Gridded maps of reconstructed tree cover through time
- S9: Data by biogeographical region
- 25 • S10: Median reconstructed tree cover based on bin-widths from other reconstructions
- S11: Median reconstructed tree cover on upper and lower age model ranges
- S12: Median tree cover reconstructions based on grid cell averages
- S13: Median tree cover reconstructions based on the Atlantic, Boreal, Continental and
Mediterranean biogeographical regions only
- 30 • S14: Supplementary information references

S1: Species categorisation

Supplementary Table 1 lists the higher level groupings of taxa considered native to Europe and included within the Total Terrestrial Pollen Sum. Species are grouped into trees, shrubs and woody vines, and herbaceous pollen (herbs, grasses etc.). Tree species are sub-divided into broadleaf and needleleaf. This data in csv form is downloadable from <https://doi.org/10.5281/zenodo.11220915>

Supplementary Table 1: Classification of taxa included within the Total Terrestrial Pollen Sum

Tree/Shrub/Non-arboreal		Taxa grouping
Tree	Broadleaf	<i>Acer, Aesculus, Alnus, Alnus subg. Alnobetula, Anacardiaceae, Arbutus, Arecaceae, Betula, Betula chamaebetula, Buxus, Carpinus, Carpinus betulus, Carpinus orientalis/Ostrya, Castanea, Celtis, Ceratonia, Cercis, Citrus, Clusiaceae, Cornaceae, Cornus, Corylus, Elaeagnaceae, Elaeagnus, Fabaceae, Fagus, Ficus, Fontanesia, Fraxinus, Ilex, Juglandaceae, Juglans, Laburnum, Laurus, Malus, Malus/Pyrus, Moraceae, Moraceae/Urticaceae, Myricaceae, Myrtaceae, Olea, Oleaceae, Pistacia, Platanus, Populus, Prunus, Pyrus, Quercus, Quercus deciduous, Quercus evergreen, Salix, Sapindaceae, Sorbus, Staphyleaceae, Syringa, Tamarix, Thymelaeaceae, Tilia, Ulmaceae, Ulmus, Ulmus/Zelkova, Verbenaceae, Viburnaceae, Viburnum, Vitex</i>
	Needleleaf	<i>Abies, Cedrus, Cupressaceae, Cupressaceae/Taxaceae, Larix, Larix/Pseudotsuga, Picea, Pinus, Pinus (diploxyylon), Pinus (haploxyylon), Taxus</i>
Shrub and woody vine		<i>Acanthaceae, Alhagi, Andromeda, Apiaceae, Araliaceae, Aristolochiaceae, Artemisia, Asparagaceae, Astragalus, Berberidaceae, Berberis, Buddleia, Calicotome, Calluna, Capparaceae, Caprifoliaceae, Cassiope, Chamaedaphne, Cistaceae, Cistus, Citrullus, Clematis, Clethra, Colutea, Convolvulaceae, Coriaria, Cotinus, Cotoneaster, Crataegus, Cucurbitaceae, Daphne, Diapensia, Dioscorea, Dioscoreaceae, Dryas, Empetrum, Ephedra, Ephedraceae, Erica, Ericaceae, Euonymus, Flueggea, Forsythia, Frangula, Frankeniaceae, Genisteae, Halimium, Hedera, Helianthemum, Hippophae, Kalmia, Lavandula, Ledum, Ligustrum, Linnaea, Lonicera, Malvaceae, Moltkia, Montiaceae, Myrica, Myricaria, Nerium, Nitrariaceae, Ononis, Osmanthus, Paeonia, Palmaria, Periploca, Phillyrea, Phlomis, Potentilla, Rhamnaceae, Rhamnus, Rhododendron, Rhus, Ribes, Rosaceae, Rubus, Ruscus, Salvia, Sambucus, Sapotaceae, Smilax, Styphnolobium, Suaeda, Vaccinium, Ziziphus</i>
Herbaceous		<i>Aconitum, Actaea, Aizoaceae, Amaranthaceae, Amaryllidaceae, Apocynaceae, Aquilegia, Araceae, Asphodelaceae, Asteraceae, Asteraceae (Liguliflorae), Asteroideae, Balsaminaceae, Boraginaceae, Brassicaceae, Campanulaceae, Cannabaceae, Carduoideae, Caryophyllaceae, Celastraceae, Chimaphila, Cichorioideae, Colchicaceae, Commelinaceae, Consolida, Crassulaceae, Cyperaceae, Datisca, Delphinium, Eriocaulaceae, Euphorbiaceae, Fabaceae (herbs), Gentianaceae, Geraniaceae, Helleborus, Hypericaceae, Impatiens, Iridaceae, Koenigia, Lamiaceae, Liliaceae, Linaceae, Linderniaceae, Linum, Lysimachia, Lythraceae, Malva, Melanthiaceae, Mercurialis, Nartheciaceae, Nigella, Onagraceae, Orchidaceae, Oxalidaceae, Oxyria/Rumex, Papaveraceae, Penthoraceae, Phyllanthaceae, Plantaginaceae, Plumbaginaceae, Poaceae, Polemoniaceae, Polygalaceae, Polygonaceae, Polygonum, Polypodiales, Portulacaceae, Primulaceae, Ranunculaceae, Ranunculus, Resedaceae, Rubiaceae, Rutaceae, Sanguisorba, Santalaceae, Saxifragaceae, Scrophulariaceae, Solanaceae, Teucrium, Thalictrum, Tofieldia, Trollius, Urticaceae, Violaceae, Zygophyllaceae</i>

S2: Coefficients and model fit based on lake sites only

Excluding bog sites from model fit improves the LOOCV RMSE to 0.13 (vs. 0.14), MAE to 0.10 (vs. 0.11) and squared correlation (R^2) of the predictions and observations to 0.69 (vs. 0.63). The model coefficients are shown in Supplementary Table 2. Coefficient estimates have the same direction as the model with bogs and lakes. Although the model fit improves, excluding bog sites from the model substantially reduces the applicability of the model to fossil records; lake records only constitute 35% of the total available fossil records.

Supplementary Table 2: Modern tree cover model coefficients for lake sites only

Coefficients (mean model with logit link)	Estimate	Standard Error	P Value
(Intercept)	-5.421	0.444	3.05e-34 ***
Tree pollen %	2.169	0.247	1.55e-18 ***
Shrub pollen %	-4.613	0.759	1.25e-09 ***
Needle share of AP%	-1.546	0.529	0.004 **
Needle share of AP%^2	3.241	0.583	2.69e-08 ***
AP Shannon index	5.239	0.483	2.14e-27 ***
AP Shannon index^2	-1.421	0.146	1.65e-22 ***
Elevation	0.001	0.001	0.317
AP pollen:elevation interaction	-0.001	0.001	0.016 *
SP pollen:elevation interaction	0.005	0.002	0.003 **
AP Shannon:elevation interaction	-0.003	0.001	0.005 **
AP Shannon^2:elevation interaction	0.001	0.000	1.02e-04 ***
Precision submodel (log link; after variable selection[^])			
(Intercept)	1.089	0.257	2.26e-05 ***
Needle share of AP%	0.751	0.271	0.006 **
AP Shannon index	0.759	0.132	9.28e-09 ***

Significance codes: 0 = '***'; 0.001 = '**'; 0.01 = '*'; 0.05 = '! 0.1; '' = 1

[^]Only significant covariates were included (at 5% significance)

S3: Coefficients and model fit without exclusion of sites >1000 m

Including higher elevation sites within the model reduces the LOOCV RMSE to 0.15 (vs. 0.14), MAE to 0.12 (vs. 0.11) and squared correlation (R^2) of the predictions and observations to 0.55 (vs. 0.63). The model coefficients are shown in Supplementary Table 3. Coefficient estimates have the same direction as the model that excludes higher elevation sites, with the exception of the dummy variable for lake or bog site, which becomes negative and significant (from positive and insignificant), and the interaction between elevation and whether a site was a lake or bog, which becomes positive (but insignificant (from negative and significant). The significance of the coefficients improves in several cases, most notably for elevation which becomes highly significant having been insignificant when excluding higher elevation sites. Although including high elevation records increases the spatiotemporal coverage of the tree cover reconstruction – the number of fossil records increases from 811 to 1050 - the impact on model fit is such that we have much less confidence in the reconstructions.

Supplementary Table 3: Modern tree cover model coefficients without exclusion of high (>1000 m) sites

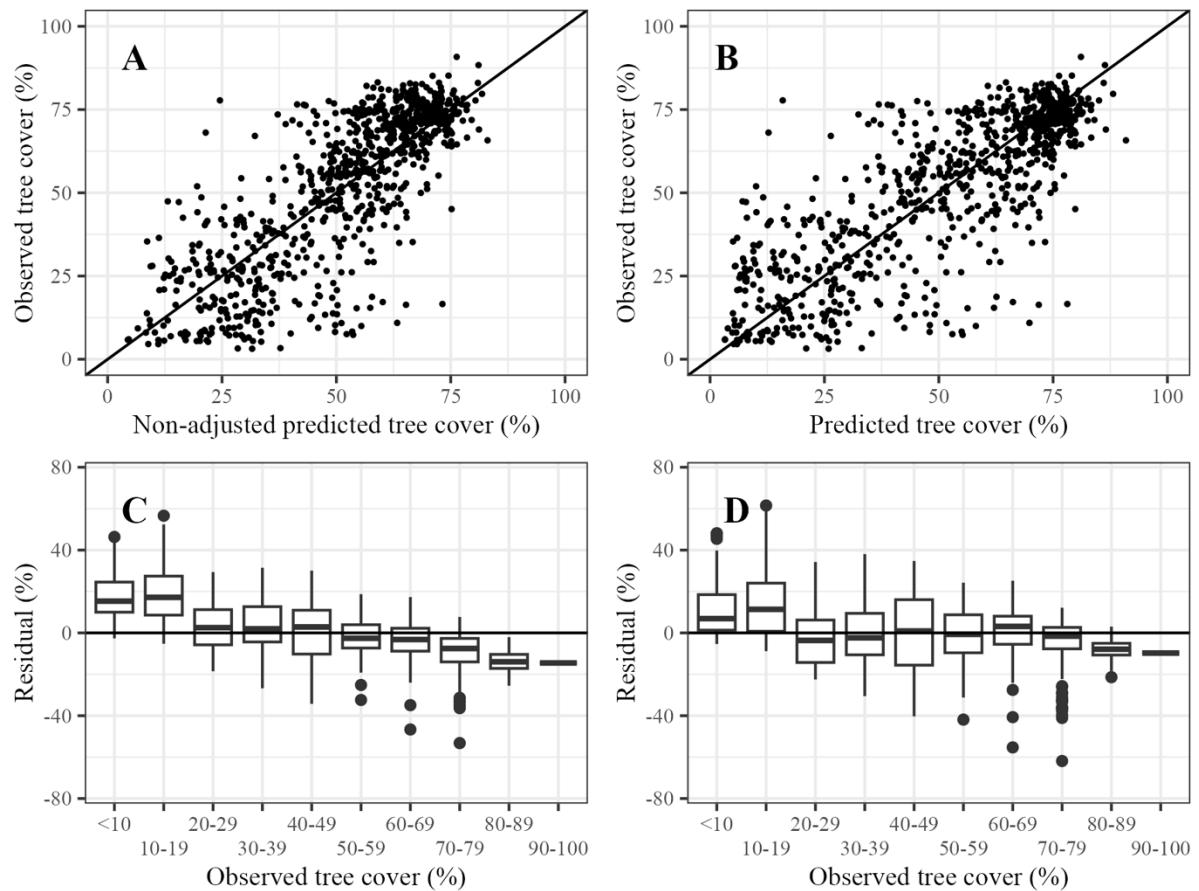
Coefficients (mean model with logit link)	Estimate	Standard Error	P Value
(Intercept)	-5.249	0.346	6.26e-52 ***
Tree pollen %	2.447	0.178	6.58e-43 ***
Shrub pollen %	-3.044	0.495	7.92e-10 ***
Needle share of AP%	-1.599	0.418	1.28e-4 ***
Needle share of AP%^2	2.954	0.453	6.98e-11 ***
AP Shannon index	4.797	0.382	4.27e-36 ***
AP Shannon index^2	-1.214	0.117	2.39e-25 ***
Lake or bog site	-0.349	0.092	1.50e-4 ***
Elevation	0.002	3.74e-4	2.71e-6 ***
AP pollen:elevation interaction	-0.001	1.83e-4	1.76e-6 ***
SP pollen:elevation interaction	0.003	5.14e-4	4.62e-10 ***
AP Shannon:elevation interaction	-0.003	1.39e-4	3.59e-10 ***
AP Shannon^2:elevation interaction	0.003	5.14e-4	1.79e-11 ***
Lake or bog site:elevation interaction	0.000	8.08e-5	0.932
Precision submodel (log link; after variable selection^)			
(Intercept)	1.146	0.228	4.81e-7 ***
Needle share of AP%	0.500	0.271	0.009 **
AP Shannon index	0.473	0.132	4.42e-6 ***
Lake or bog site	0.210	0.098	0.032 *

Significance codes: 0 = '***'; 0.001 = '**'; 0.01 = *'; 0.05 = .' 0.1; '' = 1

[^]Only significant covariates were included (at 5% significance)

S4: Quantile mapping influence on model predictions

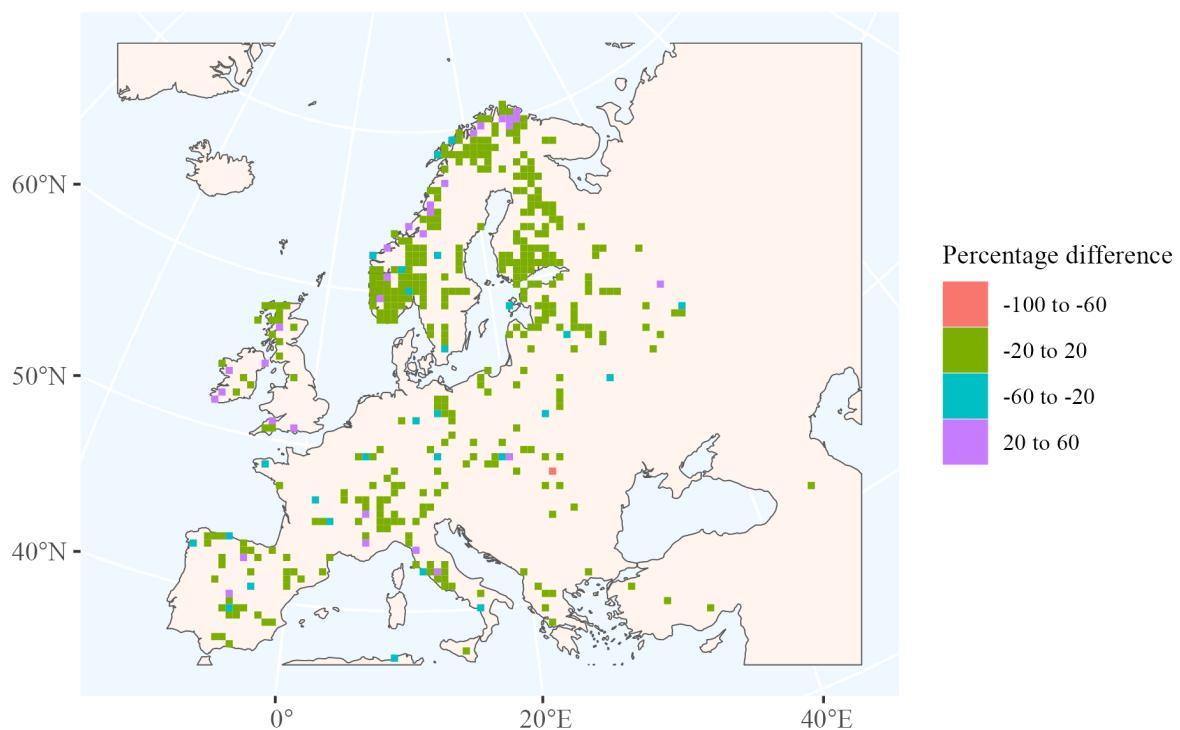
Supplementary Figure 1 (A-D) shows the impact of applying a quantile mapping model to the modern tree cover predictions solely based on the model of modern tree cover. The quantile mapping approach is designed to identify a transformation that matches the distribution of adjusted predictions to that of the observations (Gudmundsson et al., 2012). We used a smoothing spline regression curve to model the transformation, with the resulting calibration curve used to adjust modern and fossil pollen tree cover reconstructions, based on the approach from the R package *qmap* (Gudmundsson et al., 2012) and following Zanon et al. (2018). At the lower and higher levels of observed cover, quantile mapping reduces the respective over- and under-estimation.



Supplementary Figure 1: Model performance: A – Non-adjusted predictions of tree cover compared to observed tree cover; B - Predictions of tree cover compared with observed tree cover; C - Differences between non-adjusted predictions and observations (residual), in bins of observed tree cover percentage; D - Differences between predictions and observations (residual), in bins of observed tree cover percentage.

S5: Spatial structure of model predictions compared to observations

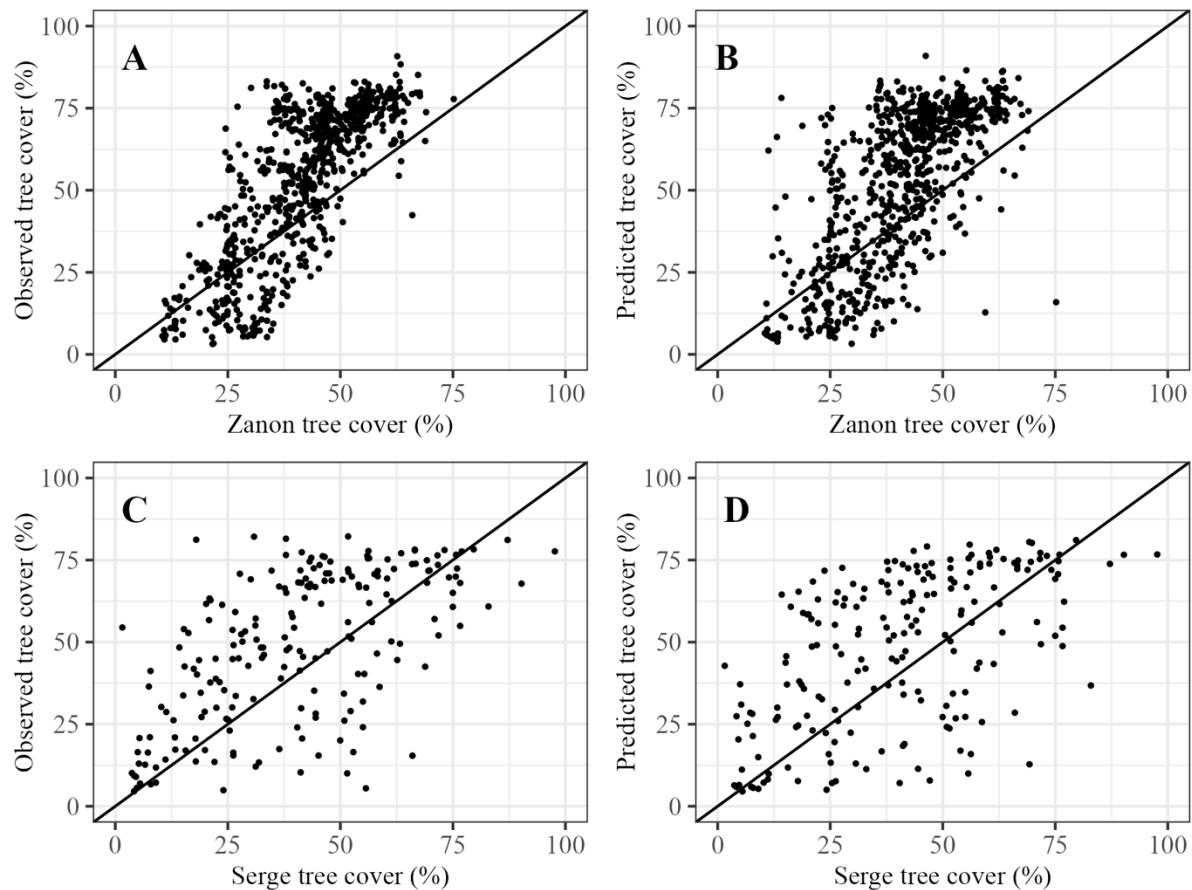
Supplementary Figure 2 is a 50km^2 map showing the grid cell averaged differences between the model predictions adjusted by quantile mapping, and observations. Difference refers to predictions minus observations. In general, there is no spatial patterning in the location of overestimates, except that there is a tendency for a grouping of cells with overestimates in the far north of Scandinavia and some sites in Ireland.



Supplementary Figure 2: Spatial structure of differences between predictions and observations.

S6: Modern predictions compared with other reconstructions based on grid cell averages

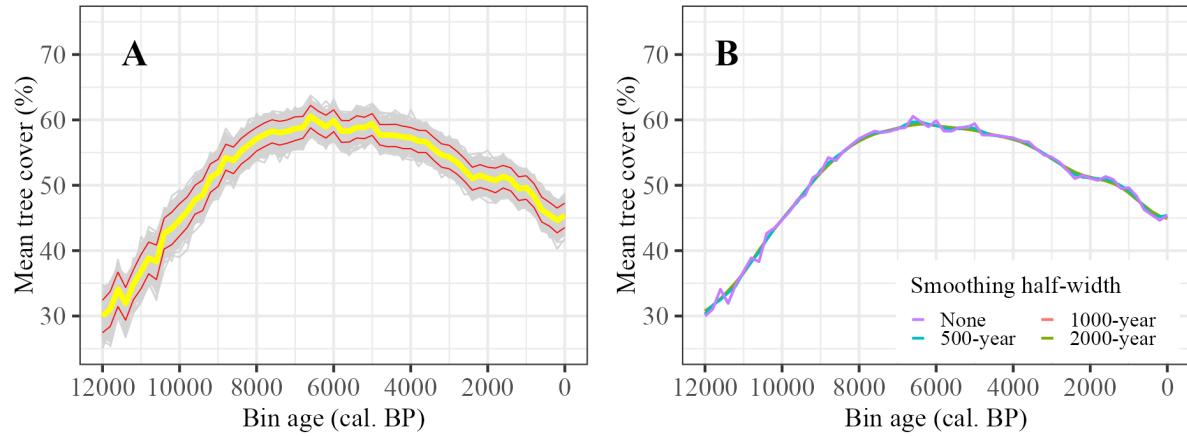
Supplementary Figure 3 (A-D) shows the relationship between observations of tree cover and our predictions of tree cover compared with values extracted from the first bin of Zanon et al. (2018) and Serge et al. (2023) at the same spatial locations. Compared with Fig. 4, Supp. Fig. 3 averages tree cover values for each record by grid cell of the other reconstructions. Hence a single point in each panel may represent tree cover value for multiple records, if there were multiple records sharing a grid cell. There are less points within panels C and D compared to A and B because Serge et al. (2023) data is provided at much lower spatial resolution than Zanon et al. (2018) data.



Supplementary Figure 3: Modern tree cover from Zanon et al (2018) per grid cell compared to (A) observed tree cover and (B) our predicted tree cover. Modern tree cover from Serge et al. (2023) per grid cell compared to (C) observed tree cover values and (D) our predicted tree cover.

S7: Mean reconstructed tree cover

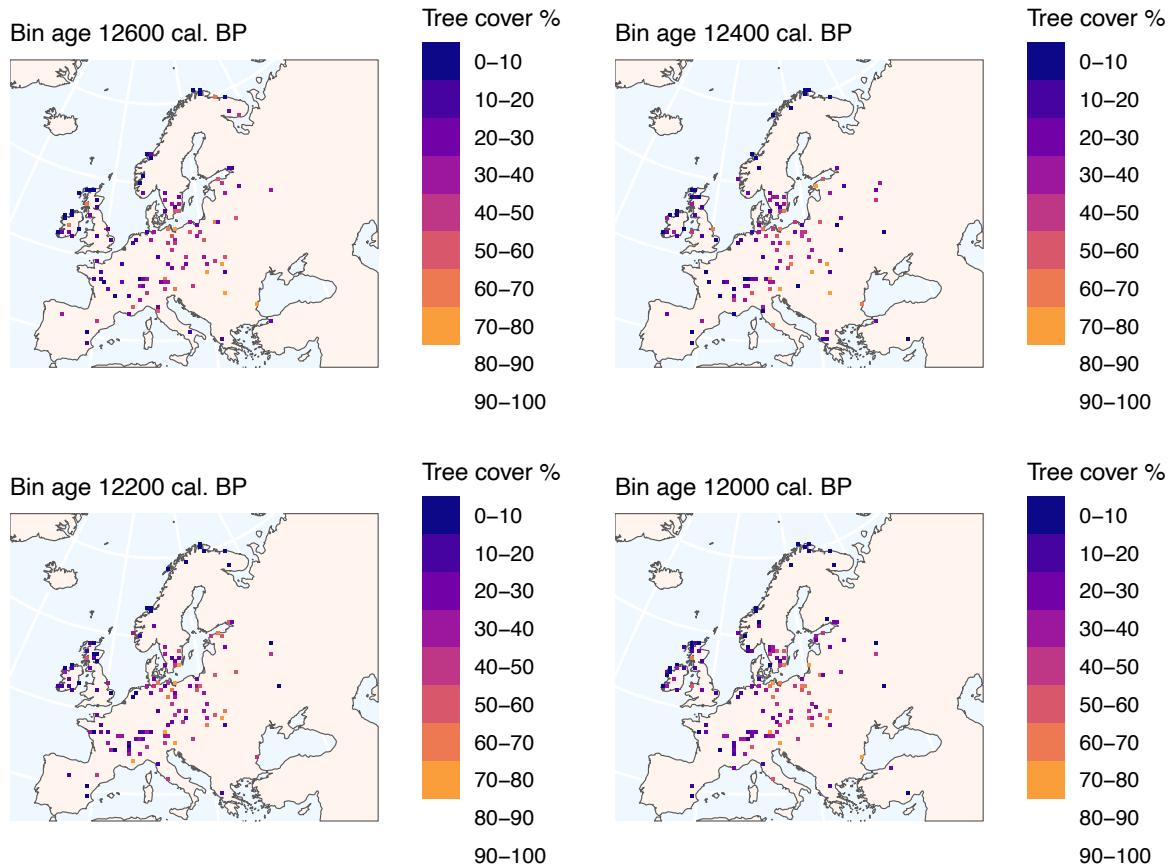
The influence of using mean tree cover values rather than median values to describe tree cover changes through time is shown in Supp. Fig. 4. Compared to the median (Fig. 5), the maximum and minimum tree cover values are reduced, and the variation through time is more limited. However, the structure of the trend through time in tree cover remains the same.

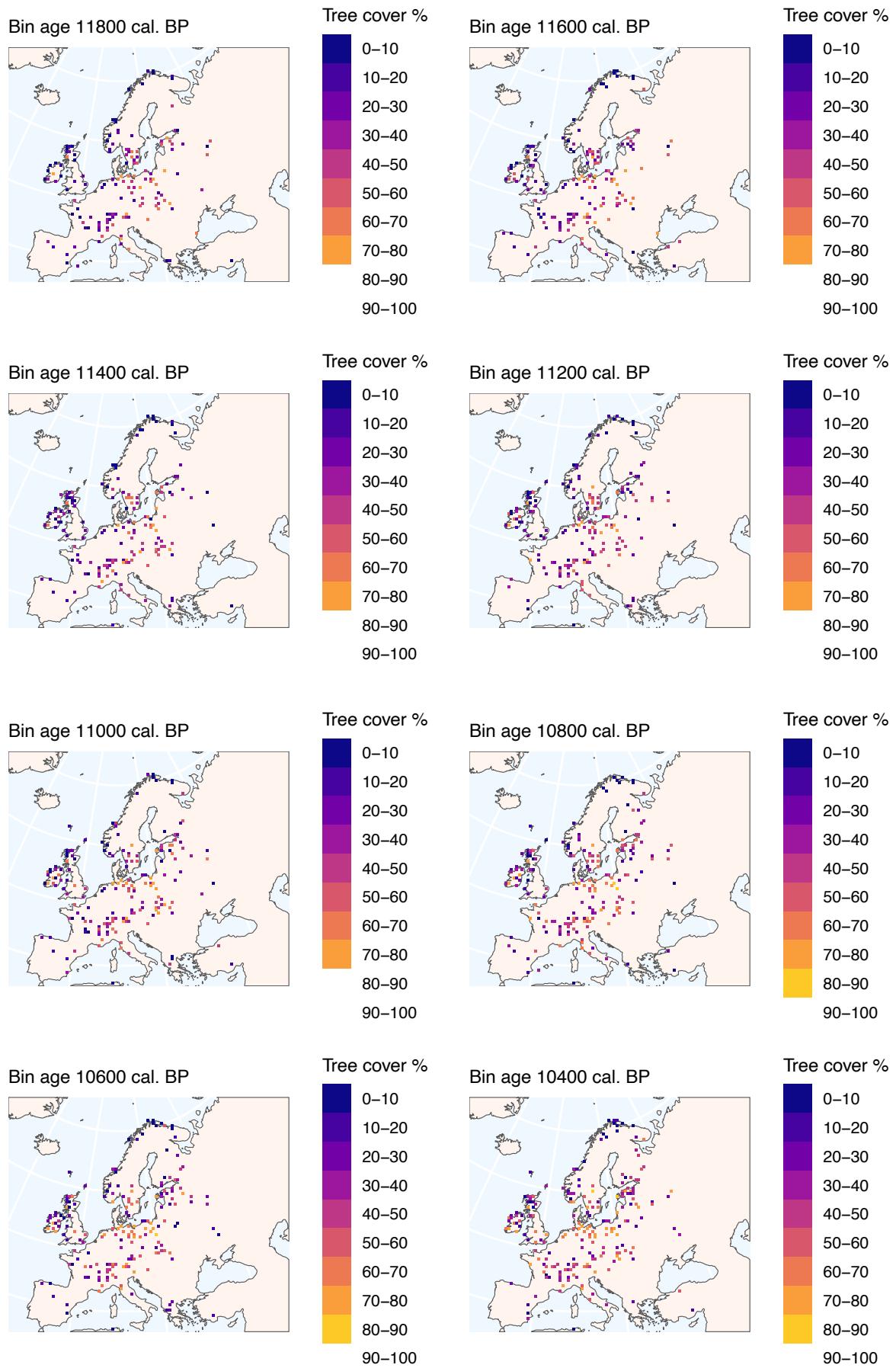


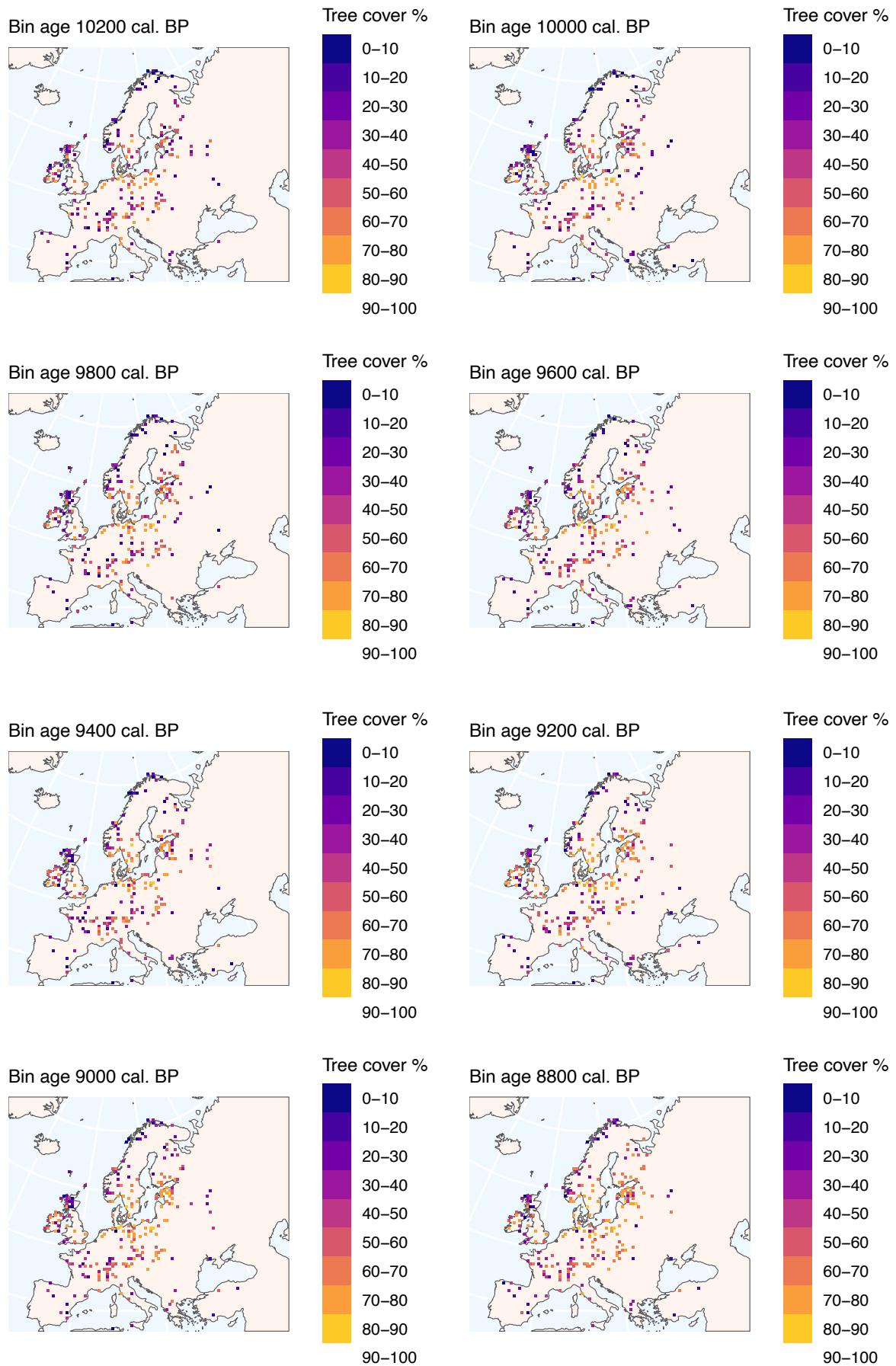
Supplementary Figure 4: A - Mean reconstructed tree cover for Europe from 12,000 to 0 cal. BP, with 95% confidence intervals for 1000 bootstrap resampling of records; B - Mean reconstructed tree cover for Europe from 12,000 to 0 cal. BP, with differing LOESS regression smoothing half-widths.

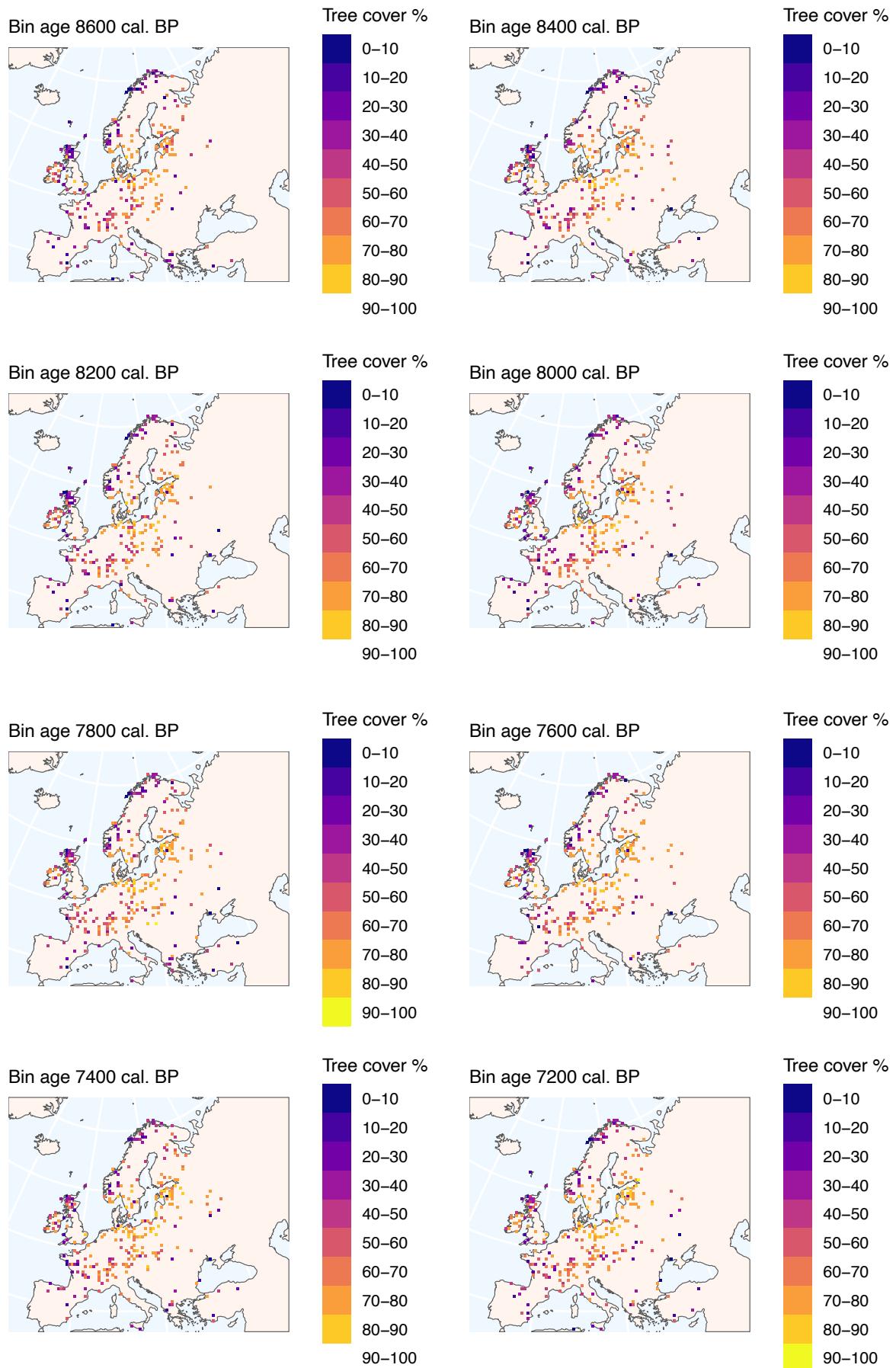
S8: Gridded maps of reconstructed tree cover through time

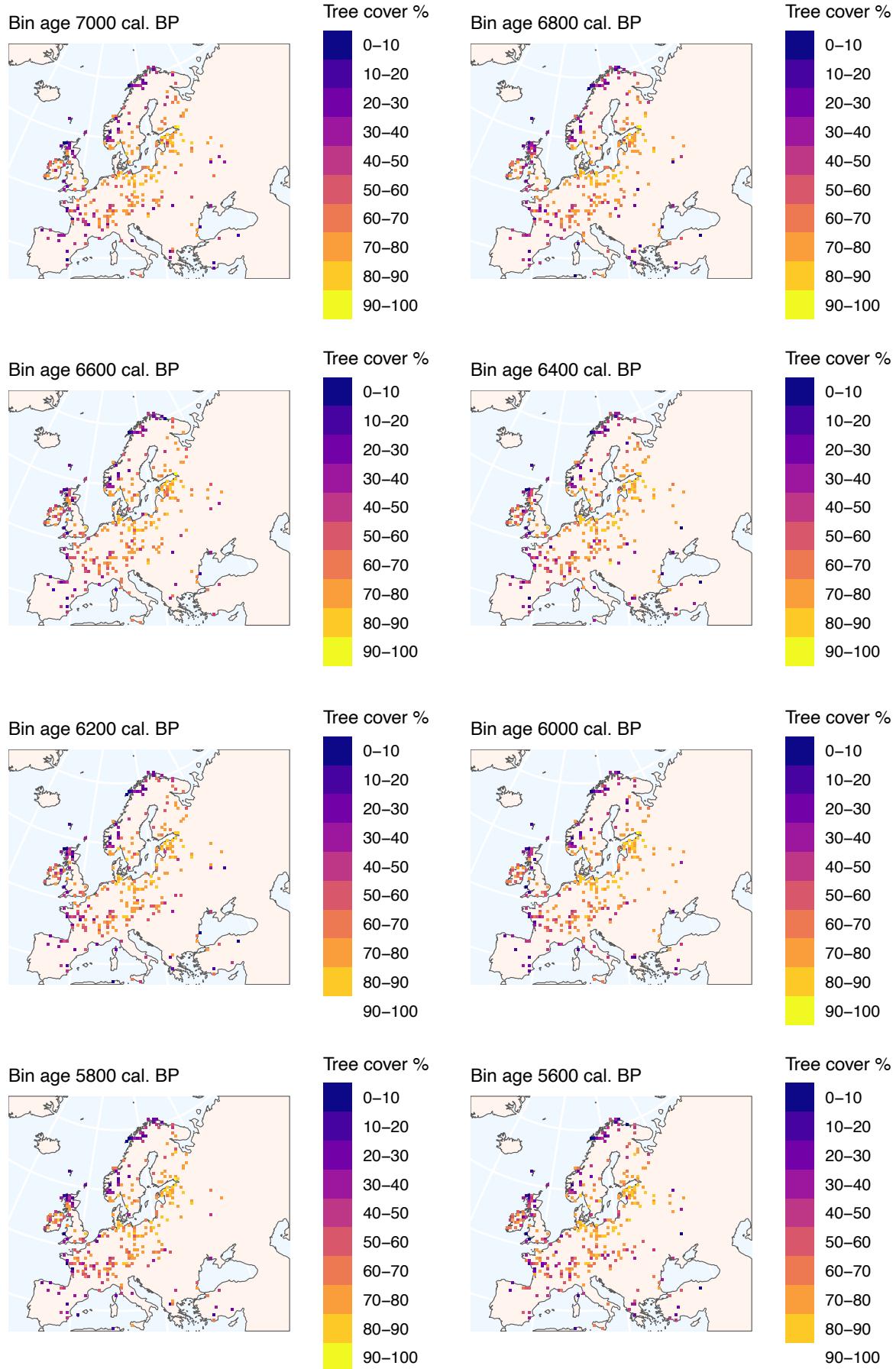
Gridded maps of reconstructed tree cover are shown below. The values in each cell are the mean of binned tree cover reconstructions for records located in each 50km² grid cell. Bins are 200-years in width, with ages referring to mid-point of each bin (e.g. the bin labelled 12,600 cal. BP represents the interval 12,700 – 12,500 cal. BP).

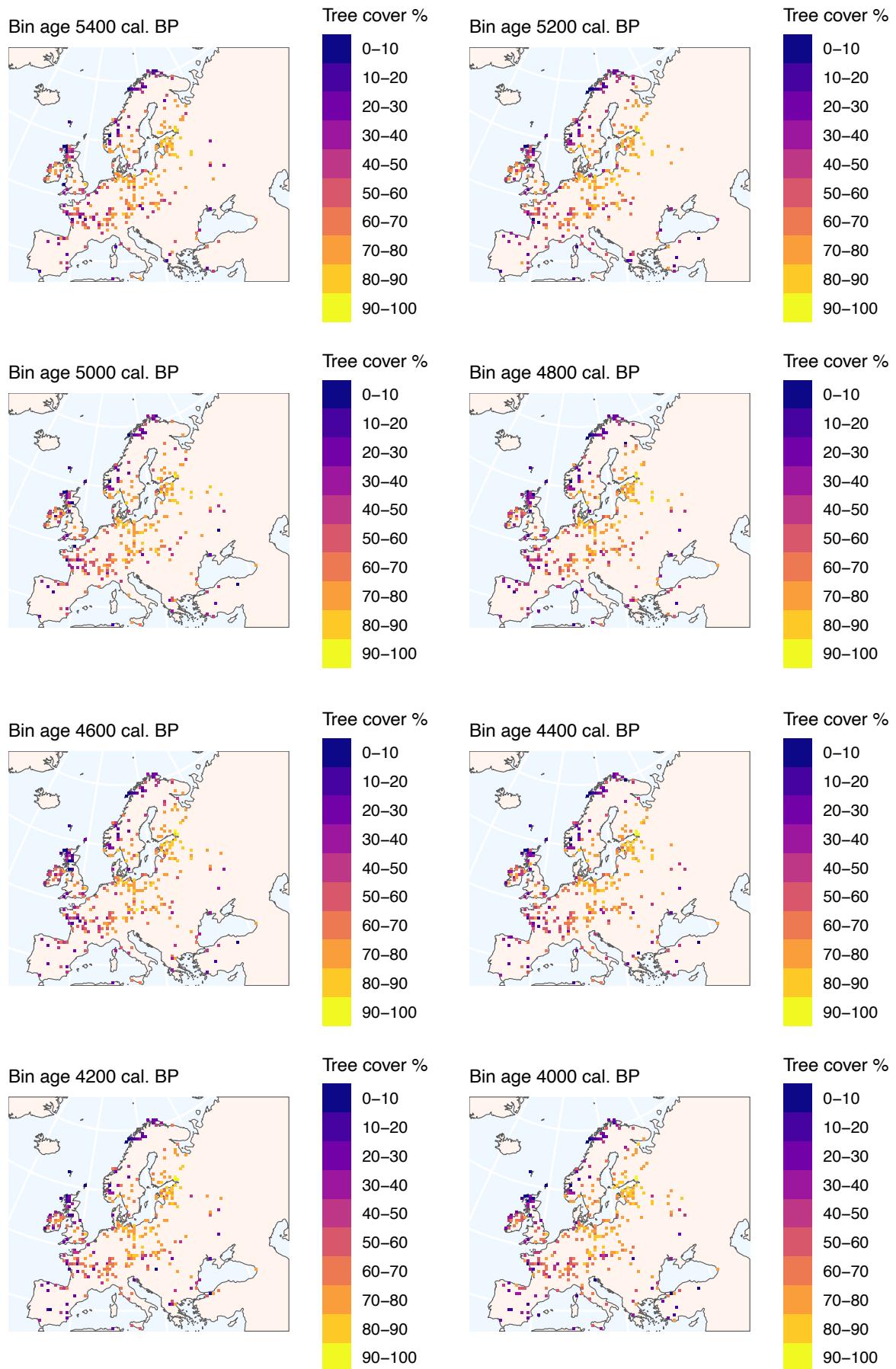


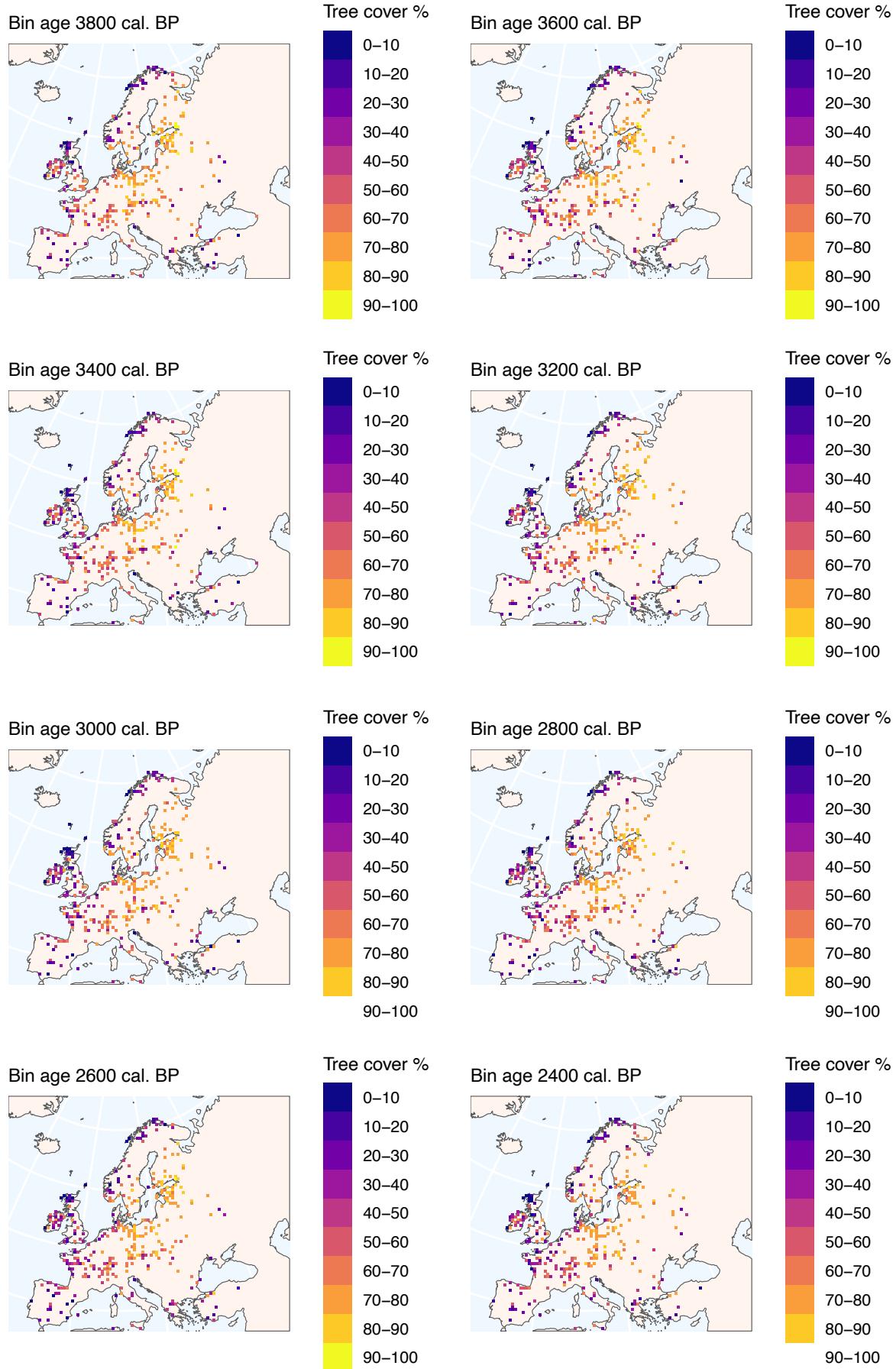


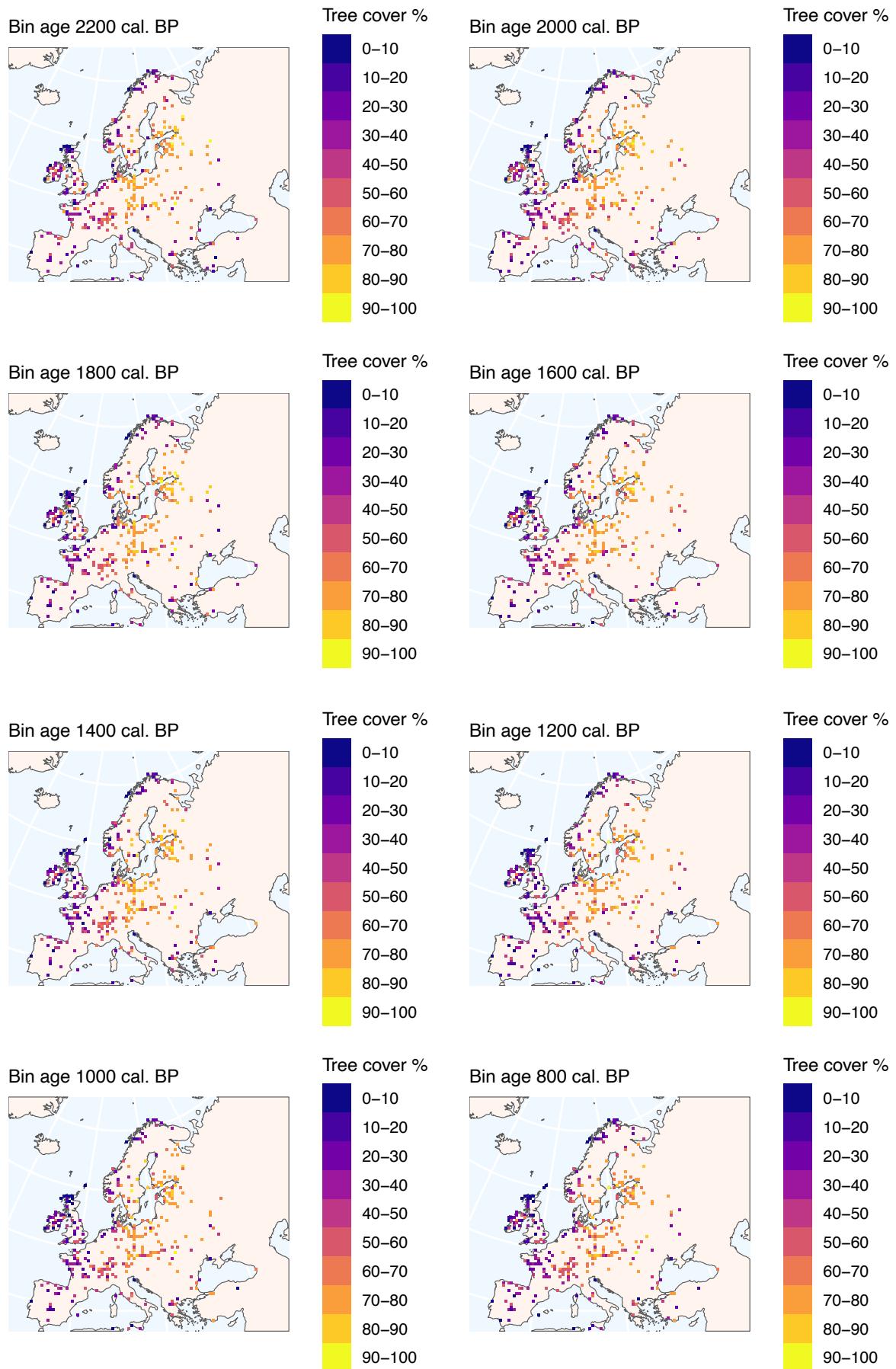


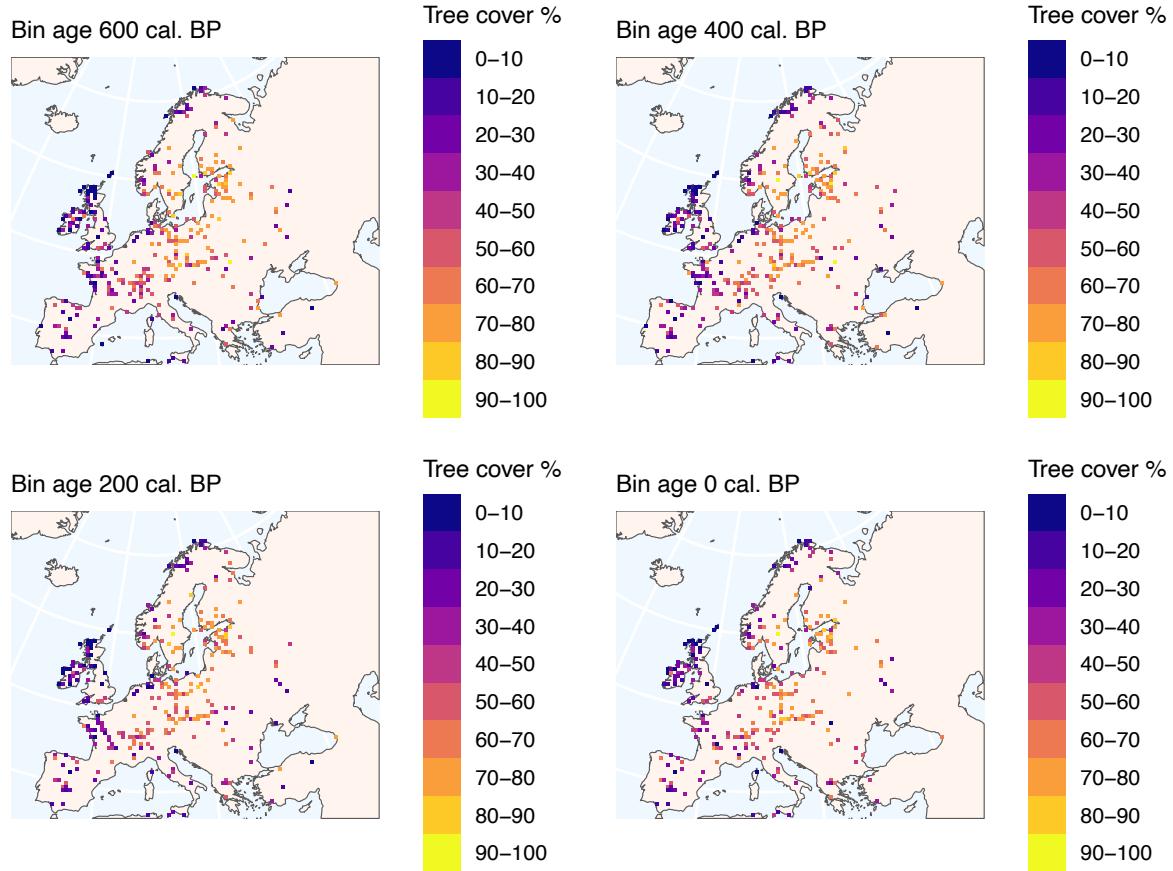












S9: Data by biogeographical region

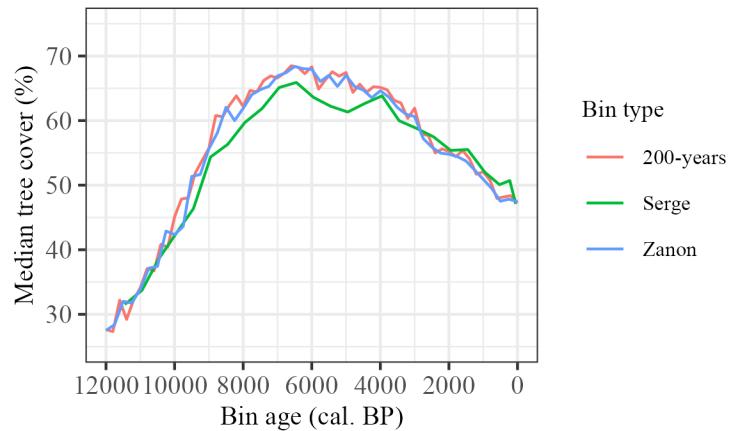
Supplementary Table 4 shows the breakdown of fossil pollen data by biogeographical region, with the number of records and the number of samples by 200-year bin within each region. Biogeographical regions are based on the European Environment Agency modern classification (European Environment Agency, 2016).

Supplementary Table 4: Data by biogeographical region

Bioregion	Number of records	Number of temporal bins
Alpine	67	2377
Anatolian	3	75
Arctic	7	279
Atlantic	284	7117
BlackSea	6	157
Boreal	137	4127
Continental	219	6840
Mediterranean	73	1701
Pannonic	13	368
Steppic	1	38

S10: Median reconstructed tree cover based on bin widths from other reconstructions

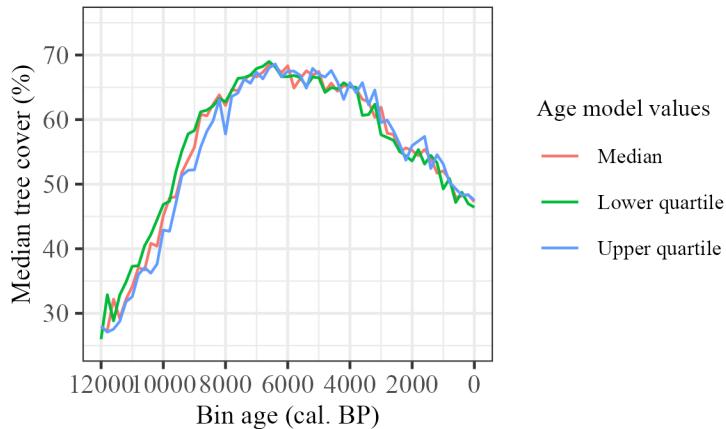
The different tree cover reconstructions use different temporal bins, which can have an influence on our reconstructed median tree cover. Supplementary Figure 5 shows our median tree cover reconstruction based on a standard 200-year bin, that same reconstruction but based on the bins used by Serge et. al. (2023) and Zanon et al. (2018). Serge et al. (2023) temporal bins cover 500-year intervals before 700 cal. BP and for 700-350 cal. BP, 350-100 cal. BP and 100 cal. BP- present. Zanon et al. (2018) temporal bins cover 250-year intervals. Although the trend in median tree cover remains the same through time, utilising the longer bin widths from Serge et al. (2023) has the effect of reducing both the variability and the range in median tree cover values.



Supplementary Figure 5: Median reconstructed tree cover for Europe from 12,000 to 0 cal. BP, based on different bin widths used in other vegetation reconstructions.

S11: Median reconstructed tree cover on upper and lower age model ranges

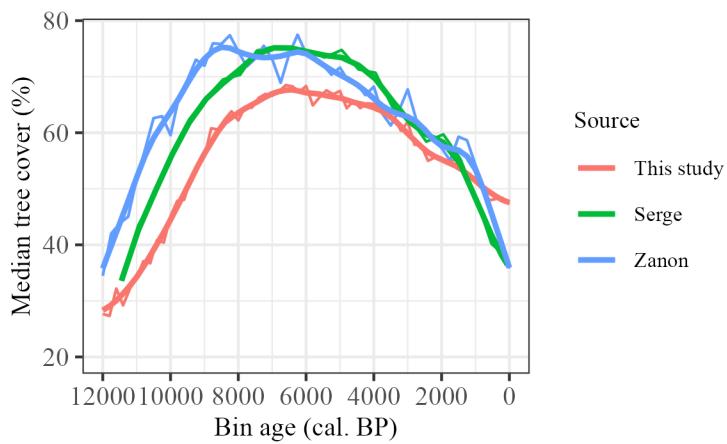
There is uncertainty associated with the dates used to construct age-depth models for the pollen records, and this is promulgated into uncertainties assigned to the sample ages in the age model. We investigated the implications of this uncertainty on median tree cover through time, by recalculating median tree cover based on the interquartile ranges of age model values (see Supp. Fig. 6). The lower quartile median value represents the 25% lower estimate of ages for each sample (older), with the upper quartile representing the 75% upper estimate of ages (younger). Due to the binning approach however, the impact of using different age model values for each sample is minimal. As Zanon et al. (2018) was published prior to the development of INTCAL20 (Reimer et al., 2020) and Marine20 (Heaton et al., 2020) calibration curves, some differences between our tree cover reconstructions may reflect differences in age models. Similarly, Serge et al. (2023) use the author's original age models, or those published in earlier datasets. Again, this use of different age models may have an impact on the comparison between median tree cover reconstructions.



Supplementary Figure 6: Mean reconstructed tree cover for Europe from 12,000 to 0 cal. BP, calculated using different fossil sample age model estimates

S12: Median tree cover reconstructions based on grid cell averages

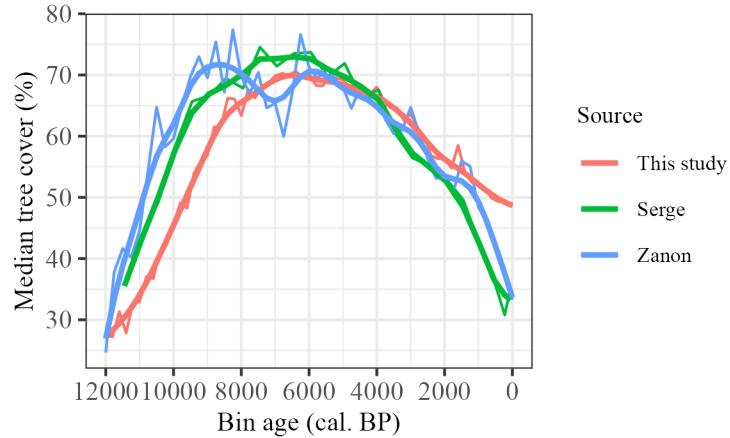
Comparisons of the median tree cover from our reconstructions with those from the Serge et al. (2023) and Zanon et al. (2018) data (Fig. 8) are based on individual pollen record locations. However, as the data from the other reconstructions are grid cell averages, the calculation of the median tree cover through time will include multiple instances of individual grid cells, where record locations share a single grid cell. This is more likely for the Serge et al. (2023) data, where reconstructions are based on larger 1° grid cells. To test the potential implications of multiple fossil record locations being within a single grid cell, Supplementary Figure 7 shows median tree cover values for these other reconstructions, but based on single values for each overlapping grid cell. Although the general trend of the tree cover medians remains the same, there is a slight increase in median tree cover values for the median reconstructions based on the data from Serge et al. (2023) and Zanon et al. (2018).



Supplementary Figure 7: Mean reconstructed tree cover for Europe from 12,000 to 0 cal. BP, with reconstructions for Serge et al. (2023) and Zanon et al. (2018) data based on grid cell averages rather than record values. Smoothed lines reflect LOESS fitted regression with 1000-year halfwidth.

S13: Median tree cover reconstructions based on the Atlantic, Boreal, Continental and Mediterranean biogeographical regions only

Limiting the data to records within the Atlantic, Boreal, Continental and Mediterranean regions has a limited effect on pan-European median tree cover reconstructions (Supp. Fig. 8).



Supplementary Figure 8: Mean reconstructed tree cover for Europe from 12,000 to 0 cal. BP, based on records within the Atlantic, Boreal, Continental and Mediterranean biogeographical regions only. Smoothed lines reflect LOESS fitted regression with 1000-year halfwidth.

S14: Supplementary information references

- European Environment Agency: Biogeographical regions,
<https://sdí.eea.europa.eu/catalogue/srv/api/records/11db8d14-f167-4cd5-9205-95638dfd9618>, 2016.
- Heaton, T. J., Köhler, P., Butzin, M., Bard, E., Reimer, R. W., Austin, W. E. N., Bronk Ramsey, C., Grootes, P. M., Hughen, K. A., Kromer, B., Reimer, P. J., Adkins, J., Burke, A., Cook, M. S., Olsen, J., and Skinner, L. C.: Marine20—The Marine Radiocarbon Age Calibration Curve (0–55,000 cal BP), *Radiocarbon*, 62, 779–820, <https://doi.org/10.1017/RDC.2020.68>, 2020.
- Gudmundsson, L., Bremnes, J. B., Haugen, J. E., and Engen-Skaugen, T.: Technical Note: Downscaling RCM precipitation to the station scale using statistical transformations - a comparison of methods, *Hydrol. Earth Syst. Sci.*, 16, 3383–3390, <https://doi.org/10.5194/hess-16-3383-2012>, 2012.
- Reimer, P. J., Austin, W. E. N., Bard, E., Bayliss, A., Blackwell, P. G., Bronk Ramsey, C., Butzin, M., Cheng, H., Edwards, R. L., Friedrich, M., Grootes, P. M., Guilderson, T. P., Hajdas, I., Heaton, T. J., Hogg, A. G., Hughen, K. A., Kromer, B., Manning, S. W., Muscheler, R., Palmer, J. G., Pearson, C., van der Plicht, J., Reimer, R. W., Richards, D. A., Scott, E. M., Southon, J. R., Turney, C. S. M., Wacker, L., Adolphi, F., Büntgen, U., Capano, M., Fahrni, S. M., Fogtmann-Schulz, A., Friedrich, R., Köhler, P., Kudsk, S., Miyake, F., Olsen, J., Reinig, F., Sakamoto, M., Sookdeo, A., and Talamo, S.: The IntCal20 Northern Hemisphere Radiocarbon Age Calibration Curve (0–55 cal kBP), *Radiocarbon*, 62, 1–33, <https://doi.org/10.1017/rdc.2020.41>, 2020.
- Serge, M., Mazier, F., Fyfe, R., Gaillard, M.-J., Klein, T., Lagnoux, A., Galop, D., Githumbi, E., Mindrescu, M., Nielsen, A., Trondman, A.-K., Poska, A., Sugita, S., Woodbridge, J., Abel-Schaad, D., Åkesson, C., Alenius, T., Ammann, B., Andersen, S., Anderson, R., Andrič, M., Balakauskas, L., Barnekow, L., Batalova, V., Bergman, J., Birks, H., Björkman, L., Bjune, A., Borisova, O., Broothaerts, N., Carrion, J., Caseldine, C., Christiansen, J., Cui, Q., Currás, A., Czerwiński, S., David, R., Davies, A., De Jong, R., Di Rita, F., Dietre, B., Dörfler, W., Doyen, E., Edwards, K., Ejarque, A., Endtmann, E., Etienne, D., Faure, E., Feeser, I., Feurdean, A., Fischer, E., Fletcher, W., Franco-Múgica, F., Fredh, E., Froyd, C., Garcés-Pastor, S., García-Moreiras, I., Gauthier, E., Gil-Romera, G., González-Sampériz, P., Grant, M., Grindel, R., Haas, J., Hannon, G., Heather, A.-J., Heikkilä, M., Hjelle, K., Jahns, S., Jasiunas, N., Jiménez-Moreno, G., Jouffroy-Bapicot, I., Kabailienė, M., Kamerling, I., Kangur, M., Karpińska-Kołaczek, M., Kasianova, A., Kołaczek, P., Lagerås, P., Latalowa, M., Lechterbeck, J., Leroyer, C., Leydet, M., Lindbladh, M., Lisitsyna, O., López-Sáez, J.-A., Lowe, J., Luelmo-Lautenschlaeger, R., Lukanina, E., Macijauskaitė, L., Magri, D., Marguerie, D., Marquer, L., Martinez-Cortizas, A., Mehl, I., Mesa-Fernández, J., Mighall, T., Miola, A., Miras, Y., Morales-Molino, C., et al.: Testing the Effect of Relative Pollen Productivity on the REVEALS Model: A Validated Reconstruction of Europe-Wide Holocene Vegetation, *Land* (Basel), 12, 986, <https://doi.org/10.3390/land12050986>, 2023.

Zanon, M., Davis, B. A. S., Marquer, L., Brewer, S., and Kaplan, J. O.: European forest cover during the past 12,000 years: A palynological reconstruction based on modern analogs and remote sensing, *Front. Plant Sci.*, 9, 1–25, <https://doi.org/10.3389/fpls.2018.00253>, 2018.