

Supporting Information

Blois et al. 10.1073/pnas.1220228110

SI Materials and Methods

Influence of Collinearity. Blois et al. (1) demonstrated that the set of variables with minimal collinearity for all times from 21 kyBP to the present (i.e., $r < 0.75$ for all time slices) consisted of summer temperature, summer precipitation, temperature seasonality, and precipitation seasonality. However, generalized dissimilarity modeling (GDM) could not fit models of compositional turnover across space for some time periods using only these four predictor variables, so Blois et al. (1) relied on a six-variable dataset that also included winter temperature and winter precipitation. In the primary analyses in this paper, we relied on the same six-variable dataset, plus CO₂, to fit models, for comparability with ref. 1. We tested whether the presence of winter temperature and winter precipitation influenced results by refitting models across space and through time [for the Late Quaternary (21–0 kyBP) dataset only] with four climate variables (summer temperature, summer precipitation, temperature seasonality, and precipitation seasonality) plus CO₂, then making predictions of spatial and temporal dissimilarity using the models (Tables S1–S3 and Fig. S4).

Influence of Geographic and Temporal Distance. We did not use geographic or temporal distance as predictors in the models, to focus on the effect of environment and because these variables are not predictive across datasets (e.g., the influence of geographic distance is 0 in the temporal models and the influence of temporal distance is 0 in the spatial models). However, we examined the influence of geographic and temporal distance on the fitted functions for the primary datasets (e.g., the spatial and temporal datasets from the entire period 21–0 kyBP) by fitting models with and without distance as a predictor, in addition to the seven environmental variables (Fig. S5). Geographic distance was modeled using latitude and longitude, with latitude and longitude projected to North America Albers equal-area conic projection (origin: 40°N, 96°W; standard parallels: 20°N, 60°N). Temporal distance was modeled with years before present as the x -value (e.g., 1, 2, 3 kyBP) and 0 as the y -value.

Community Subsampling. The focus of the paper is on comparing and processing datasets based on characteristics of the climate, but there also are differences between the distribution of spatial and temporal compositional dissimilarities (Fig. S6). The overall range of community dissimilarity is similar between spatial and temporal datasets, but the shape of the dissimilarity distribution is different: the temporal models are particularly dominated by low dissimilarity between sites, which inherently is more difficult to model (Fig. S6). We tested whether we can improve space-for-time substitution by preprocessing datasets so that the community dissimilarities are similar, similar to the approach taken for preprocessing the datasets based on climate dissimilarity. Essentially, we are asking, Can we build substitutable models if we focus only on the subset of sites with similar differences in their pollen profiles? We reprocessed all datasets, focusing only on the pairwise differences between taxa rather than climate variables. To do this, we reran *MatchIt* on the pairwise relative abundance differences between taxa, focusing only on the 10 most abundant taxa across the entire dataset (*Pinus*, *Betula*, *Quercus*, *Picea*, *Alnus*, *Tsuga*, *Artemisia*, *Ulmus*, *Ambrosia*-type, and *Fagus*). We then refit models to each training dataset and predicted spatial and temporal dissimilarity of the subsampled evaluation datasets (Tables S1–S3 and Fig. S2).

Discussion of GDM Analyses. Fit of GDM across space and through time. Environment explains a large, but differential, percentage of deviance in compositional dissimilarity across space and time (46.8% of deviance explained in the temporal model, 39.6% of deviance explained in the spatial model) (Table S3). The greater percentage of deviance explained by the temporal model probably is because many factors besides climate may influence compositional dissimilarity, and these unknown and unmodeled variables likely vary more across space (e.g., between sites) than through time (e.g., between times at the same site).

Predictions of spatial dissimilarity. Time-for-space substitution generally did not perform as well as space-for-time substitution, achieving 54% predictive skill for the Late Quaternary (Fig. S3 and Table S2). For the full dataset (21–0 kyBP), goodness of fit between predictions of spatial turnover from the spatial and temporal models was reduced major axis (RMA) $R^2 = 0.5764$ (compared with RMA $R^2 = 0.7116$ for the corresponding predictions of temporal turnover) (Table S1). Time-for-space substitution during the Late Pleistocene (21–11 kyBP) also did not perform as well (model-to-model regression based on predictions of spatial dissimilarity from the spatial and temporal models: RMA $R^2 = 0.7826$, compared with RMA $R^2 = 0.8329$ for the corresponding temporal dissimilarity predictions) (Table S1). Goodness of fit between predictions of spatial dissimilarity were relatively higher in the Holocene (RMA $R^2 = 0.3606$, vs. RMA $R^2 = 0.199$ for temporal dissimilarity predictions), but were still weaker than in the Pleistocene. Goodness of fit between observed dissimilarity and spatial dissimilarity predictions varied among models (Table S2).

Thus, similar to space-for-time substitution, time-for-space substitution performs well in the Pleistocene (79%), but in the Holocene it is relatively weak (29%). However, the source of the problem is opposite that of space-for-time substitution and stems from the “no-analog” problem (2); e.g., the temporal models are calibrated on a narrow range of data (temporal climate dissimilarity in the Holocene) but used to predict compositional dissimilarity among sites with much larger climate dissimilarities.

Influence of collinearity. We found that collinearity between winter temperature or winter precipitation and the other climate variables did not substantially influence the fitted models or predictions. Predictions (Tables S1 and S2), fitted functions (Fig. S4), and deviance explained (Table S3) for the five-variable model were very similar to the original, seven-variable models.

Influence of geographic and temporal distance. Distance was a significant variable in both the spatial and temporal models. The percentage of deviance explained increased slightly, around 2%, for both models when distance was included as a variable (Table S3). The inclusion of geographic distance did not change the fitted functions for the environmental variables substantially, although winter precipitation is a significant predictor when geographic distance is not included in the spatial model (Fig. S5). Similarly, the temporal model is similar with and without temporal distance as a predictor (Fig. S5). The main difference is in temperature seasonality; this variable is significant only when temporal distance is not included in the model.

Influence of CO₂. CO₂ also influenced the models only slightly, increasing the deviance explained in the temporal model by about 1.5 percentage points (e.g., 46.8% deviance explained in the temporal model with CO₂, vs. 45.3% without CO₂; Table S3). However, we kept this variable in the model because it is an important environmental variable that has changed substantially

since preindustrial times and will continue to change in the future.

Climate subsampling. The Pleistocene model did not change much based on subsampling by climate. The fitted functions differ a bit more between the spatial and temporal models in the climate-subsampled Pleistocene models than in the full Pleistocene models (Fig. S2 C and D). Deviance explained by the spatial model was considerably less (18.6% in the climate-subsampled Pleistocene vs. 28.2% in the full Pleistocene), but the temporal model improved slightly (34.1% in the climate-subsampled Pleistocene vs. 32.2% in the full Pleistocene) (Table S3). Predictions did not change appreciably (Fig. 3 B, C, and E), and space-for-time substitution still achieved 72% predictive skill (Table S2).

The fitted functions for the climate-subsampled Holocene temporal model were essentially identical to the unsampled Holocene model (because very few sites were removed by subsampling; Fig. 3). However, the climate-subsampled Holocene spatial fitted functions were different from those fitted by the unsampled Holocene dataset (Fig. S2E), and the deviance explained by climate in the subsampled spatial model was reduced (e.g., 26.8% vs. 44.1% in the unsampled Holocene model) (Table S3). Overall, the correspondence between the spatial and temporal models improved (Fig. S2 E and F), and predictions of

temporal dissimilarity also improved (Fig. 3 G, H, and J and Tables S1 and S2).

Community subsampling. Compositional dissimilarity profiles were very similar across the subsampled spatial and temporal datasets (Fig. S6 I and L), and the climate dissimilarities that were carried along in the subsampling also were similar (Fig. S6 C and F). The resulting community-subsampled spatial and temporal models are very similar to one another in all time partitions (21–0 kyBP, 21–11 kyBP, and 10–0 kyBP) (Fig. S2). There still are differences in the fitted functions, but they are relatively minor and predictions are tight (Table S1). Note, however, that the maximum height along the y-axis attained by the fitted functions is much less than in either the original or climate-subsampled models, particularly across space, indicating that total community turnover is much less in the community-subsampled models (Fig. S2). Subsampling based on the community means we are explicitly choosing sites that have similar communities, so dissimilarity/turnover gradients necessarily are less. Additionally, goodness of fits between the predictions and observations of dissimilarity are among the weakest in any of the models we have considered, although space-for-time substitution has high predictive skill (Table S2). Thus, although subsampling based on community made the models and predictions tighter across space and time, overall it is less meaningful.

1. Blois JL, et al. (2013) Modeling the climatic drivers of spatial patterns in vegetation composition since the Last Glacial Maximum. *Ecography* 36(4):460–473.

2. Williams JW, Jackson ST (2007) Novel climates, no-analog communities, and ecological surprises. *Front Ecol Environ* 5(9):475–482.

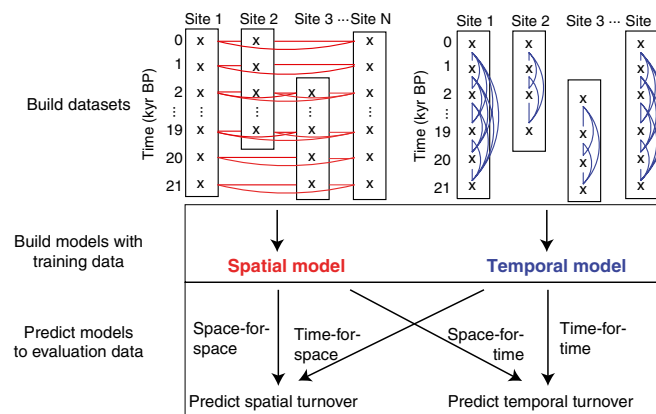


Fig. S1. Experimental design. Red indicates spatial model, blue indicates temporal model. The main article presents space-for-time substitution results, relative to time-for-time performance, whereas *SI Materials and Methods* also presents time-for-space relative to space-for-space results.

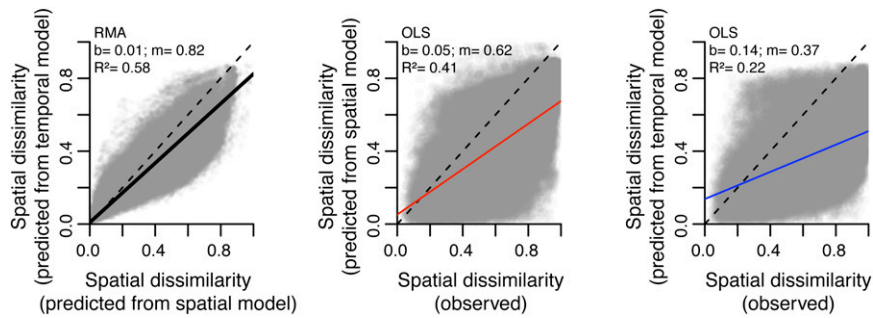


Fig. 53. Spatial dissimilarity predictions (e.g., time-for-space substitution) for the Late Quaternary dataset (21–0 kyBP). (*Left*, solid black line) RMA regression between spatial dissimilarity predicted by the spatial model and by the temporal model. (*Center*, red line) Ordinary least-squares (OLS) regression between observed spatial dissimilarity and spatial dissimilarity predicted by the spatial model. (*Right*, blue line) shows the OLS regression between observed spatial dissimilarity and spatial dissimilarity predicted by the temporal model. The dashed lines indicate the 1:1 lines.

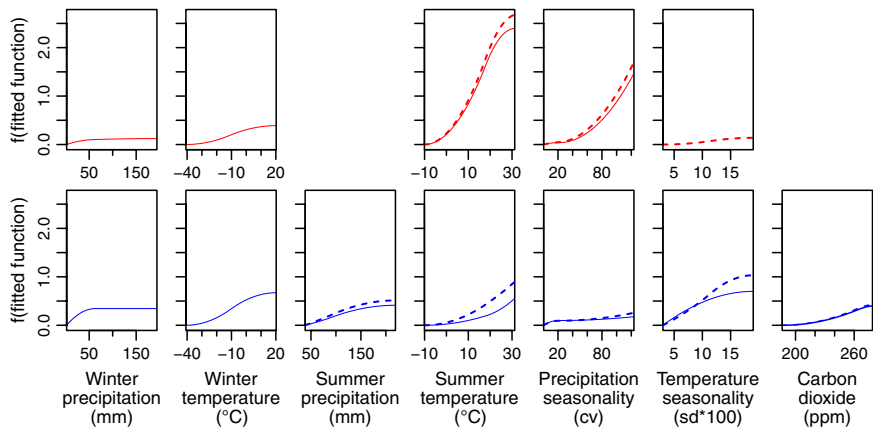


Fig. 54. Final fitted functions for spatial (*Upper*, red) and temporal (*Lower*, blue) models (Late Quaternary dataset: 21–0 kyBP) with (solid lines) and without (dashed lines) winter temperature and winter precipitation as potential predictors.

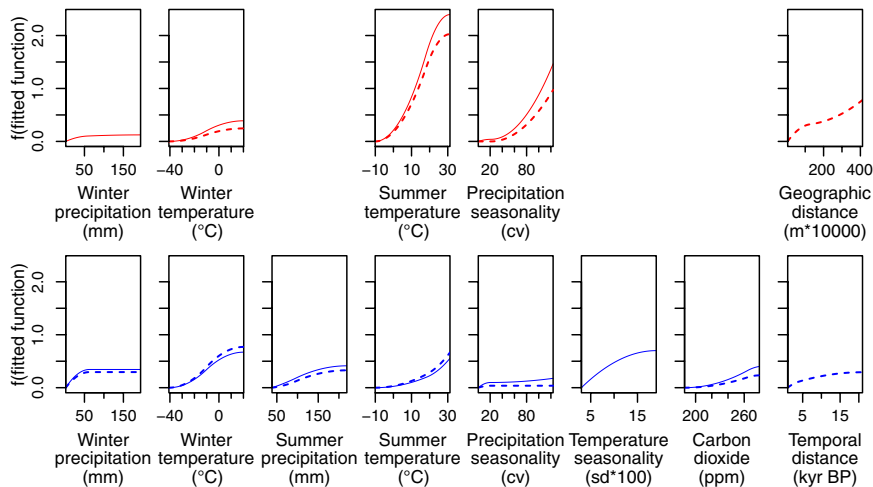


Fig. 55. Fitted functions for spatial (*Upper*, red) and temporal (*Lower*, blue) models (Late Quaternary dataset: 21–0 kyBP) with (dashed lines) and without (solid lines) distance as a predictor.

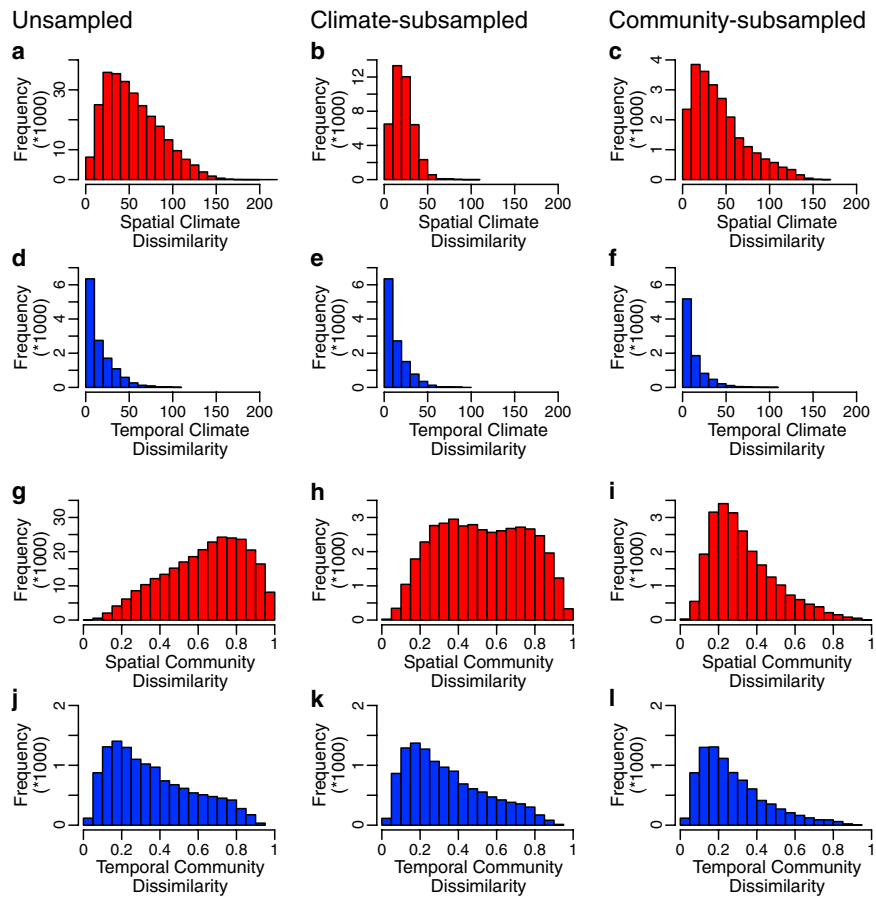


Fig. S6. Euclidean dissimilarity in climate (A–F) and Bray–Curtis dissimilarity in the community (G–L) in the unsampled (A, D, G, and J), climate-subsampled (B, E, H, and K), and community-subsampled (C, F, I, and L) datasets.

Table S1. Model-to-model RMA regression statistics for the main and supplemental analyses

Dataset	Temporal turnover			Spatial turnover		
	b	m	R^2	b	m	R^2
Late Quaternary (21–0 kyBP)	–0.08	1.1	0.7116	0.01	0.82	0.5764
Late Pleistocene (21–11 kyBP)	–0.04	0.89	0.8329	0.04	0.72	0.7826
Holocene (10–0 kyBP)	–0.05	0.91	0.199	–0.05	0.98	0.3606
Late Quaternary five-variable	–0.09	1.1	0.6799	0.01	0.77	0.5402
Late Quaternary climate-subsampled	–0.03	1.18	0.8388	–0.01	0.68	0.8012
Late Pleistocene climate-subsampled	–0.04	1.01	0.7684	0	0.74	0.7305
Holocene climate-subsampled	–0.02	1.07	0.5504	0	0.56	0.4886
Late Quaternary community-subsampled	–0.02	0.53	0.759	0.02	1.32	0.8644
Late Pleistocene community-subsampled	–0.02	0.68	0.7426	–0.03	1.66	0.4594
Holocene community-subsampled	–0.01	0.31	0.5926	–0.04	2.06	0.9239

Shown are the name of the dataset, the RMA model regression statistics for temporal turnover predicted by the spatial model and by the temporal model, and the RMA model regression statistics for spatial turnover predicted by the temporal model and by the spatial model. b, intercept of the regression line; m, slope of the regression line; R^2 , goodness of fit.

Table S2. Model-to-observation OLS regression statistics for the main and supplemental analyses

Dataset	Space-for-time results							Time-for-space results						
	Space for time			Time for time			Space-for-time skill, %	Space for space			Time for space			Time-for-space skill, %
	b	m	R ²	b	m	R ²		b	m	R ²	b	m	R ²	
Late Quaternary (21–0 kyBP)	0	0.55	0.3432	0.04	0.58	0.4773	72	0.05	0.62	0.4135	0.14	0.37	0.2222	54
Late Pleistocene (21–11 kyBP)	0.1	0.29	0.2363	0.14	0.36	0.2839	83	0.05	0.48	0.299	0.1	0.31	0.2355	79
Holocene (10–0 kyr BP)	0.08	0.08	0.0287	0.1	0.24	0.1984	14	0.04	0.67	0.4583	0.19	0.35	0.1315	29
Late Quaternary five-variable	0.01	0.52	0.3227	0.0	0.57	0.4724	68	0.05	0.62	0.4083	0.13	0.34	0.2103	52
Late Quaternary climate-subsampled	0.05	0.57	0.3749	0.05	0.53	0.4512	83	0.03	0.65	0.4849	0.03	0.4	0.3827	79
Late Pleistocene climate-subsampled	0.12	0.31	0.218	0.14	0.37	0.3044	72	0.07	0.31	0.2411	0.08	0.19	0.1721	71
Holocene climate-subsampled	0.11	0.18	0.0933	0.1	0.24	0.1985	47	0.09	0.37	0.2871	0.09	0.13	0.1112	39
Late Quaternary community-subsampled	0.02	0.17	0.2032	−0.05	0.37	0.2946	69	0.04	0.31	0.2345	0.08	0.39	0.2072	88
Late Pleistocene community-subsampled	0.07	0.14	0.0909	0.09	0.35	0.2414	38	0.09	0.1	0.0612	0.1	0.19	0.0857	140
Holocene community-subsampled	0.02	0.05	0.0923	0.07	0.22	0.1713	54	0.02	0.3	0.2304	0.01	0.63	0.236	102

Shown for each dataset are the dataset name, statistics for the space-for-time OLS regression between observed temporal turnover and temporal turnover predicted by the spatial model, statistics for the time-for-time OLS regression between observed temporal turnover and temporal turnover predicted by the temporal model, the predictive skill of space-for-time substitution (i.e., space-for-time R² divided by time-for-time R²), statistics for the space-for-space OLS regression between observed spatial turnover and spatial turnover predicted by the spatial model, statistics for the time-for-space OLS regression between observed spatial turnover and spatial turnover predicted by the temporal model, and the predictive skill of time-for-space substitution (i.e., time-for-space R² divided by space-for-space R²). Abbreviations are as in Table S1.

Table S3. Percentage of deviance explained for each model

Dataset	Domain	% deviance explained
Late Quaternary (21–0 kyBP)	Spatial	39.6
	Temporal	46.8
Late Pleistocene (21–11 kyBP)	Spatial	28.2
	Temporal	32.2
Holocene (10–0 kyBP)	Spatial	44.1
	Temporal	21.7
Late Quaternary five-variable	Spatial	39.1
	Temporal	46.3
Late Quaternary climate-subsampled	Spatial	49.9
	Temporal	43.5
Late Pleistocene climate-subsampled	Spatial	18.6
	Temporal	34.1
Holocene climate-subsampled	Spatial	26.8
	Temporal	21.6
Late Quaternary community-subsampled	Spatial	22.1
	Temporal	30.5
Late Pleistocene community-subsampled	Spatial	14.6
	Temporal	28.8
Holocene community-subsampled	Spatial	26.2
	Temporal	21.5
Late Quaternary—no CO ₂	Spatial	39.6
	Temporal	45.3
Late Quaternary environment + distance	Spatial	41.5
	Temporal	48.8

Rows correspond to those in Tables S1 and S2, with the addition of rows at the bottom for spatial and temporal models with no CO₂ and spatial and temporal models with distance added. “Domain” indicates whether the model is explaining spatial or temporal compositional dissimilarity.