Detailed methods:

Simulated Phenology Data

To simulate phenological data for locations ranging from 23˚N to 68˚N, we used the number of sites and years at each location, and the mean temperature at each location, from the dataset compiled by Huang *et al*.[1].We first simulated daily temperatures that followed a seasonal warming trend, typical of the transition from winter to spring, using the mean annual temperature (MAT) from each site and year in the following way. The "winter" in our simulated data lasted 60 days, beginning on 1 January for all sites, and daily temperatures were normally distributed, with a mean temperature 3˚C cooler than the MAT at that site and year. Following this, there were two "inter-season" periods, each for 30 days, with mean temperatures that were 2˚C cooler than and equal to MAT, respectively. Next, there was a warm season with a mean temperature of 3˚C warmer than MAT. These daily temperatures were used to calculate daily values for forcing sums (i.e., growing degree days), following the approach of Huang *et al*. (1) i.e., using the same forcing function and threshold temperature for onset, Fig. 1A).

We generated simulated phenology responses using only forcing and latitude.. We varied the forcing units required for onset of wood formation (FU) with latitude, using the following equation, derived from the fitted relationship present in the Huang *et al.* (1) dataset (Fig. 1B):

$$FU = 575.477 + -8.238 * latitude$$

We used this simple model for two reasons: First, simple models generally perform as well as more complex models at predicting phenology (2). However, phenology models that used a fixed FU threshold tend to show latitudinal biases, with more northern populations requiring less forcing (3). Similarly, in this dataset, the annual cumulative forcing in some northern populations is less than the dataset mean FU_{wf}.

We obtained photoperiod values for each simulated phenology event at the latitude and longitude for each site and year from (1), using the insol package in R. To compare the explanatory power of photoperiod to MAT, we fit separate linear regressions, across all sites, and compared r-squared values.

Predictive performance of Forcing, Latitude, MAT, and Photoperiod

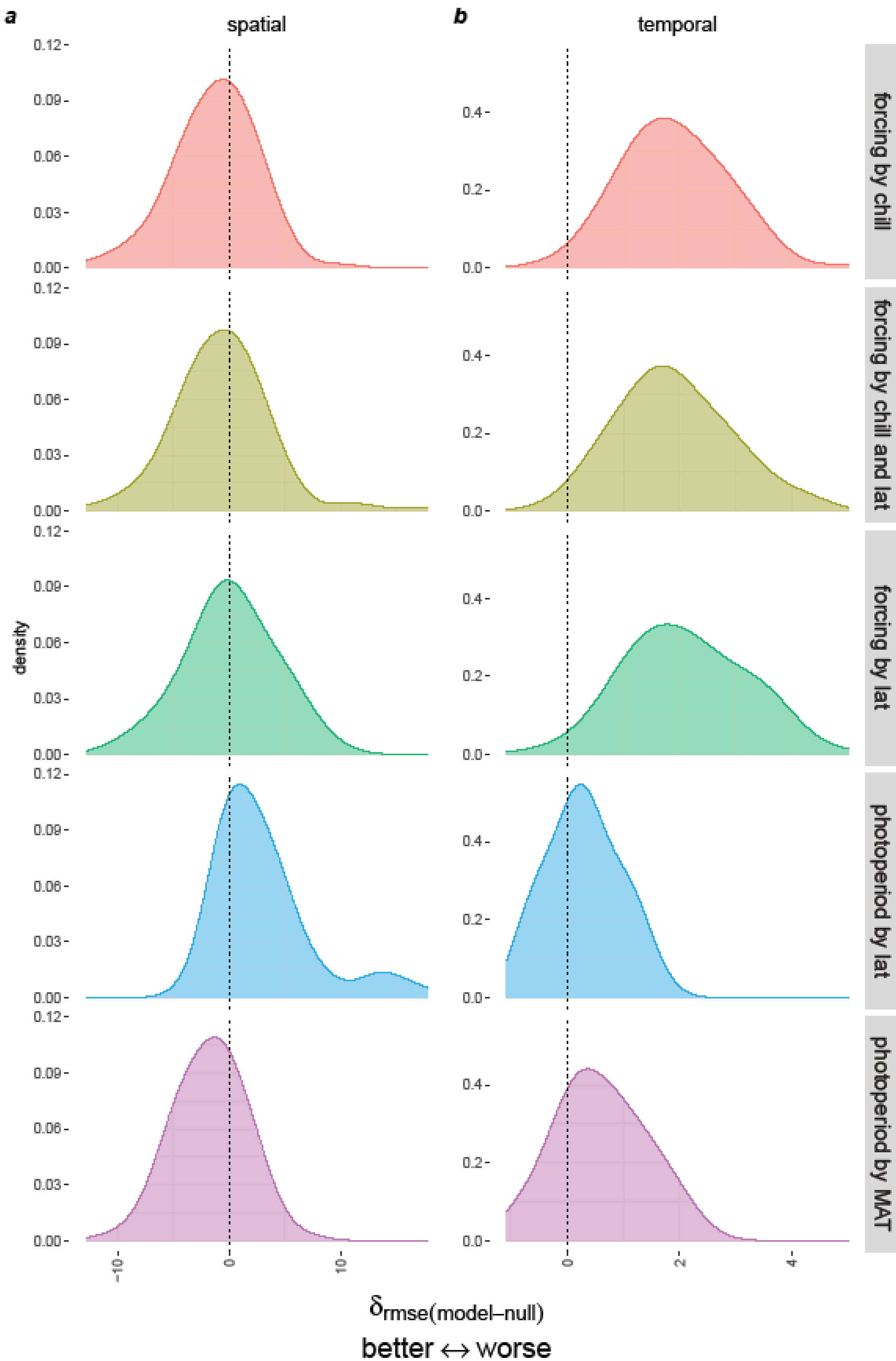
We reanalyzed the original xylogenesis phenology dataset from Huang *et al.*'s Table S1 to test the predictive performance of a suite of models (Table S1; below) that included forcing photoperiod, latitude, MAT, and chilling, as well as interactions between them (e.g. a model where forcing thresholds vary with MAT or forcing sums required vary by latitude or with chilling). For the forcing and chilling predictions, we extracted site-specific daily temperatures from the ERA 5 daily reanalysis product (4) via Google Earth Engine. To adjust for any biases that might be introduced by using a gridded meteorological product rather than a local met station, we adjusted the ERA 5 data using the difference in mean annual temperature from local station data provided by Huang *et al.* The resulting estimates are generally close to the forcing/chilling calculations using local meteorological stations provided in Huang *et al.*'s Table S1, although we acknowledge model performance for forcing and chilling models might be improved with in situ temperature data. The latitude, null, and MAT models yield a direct prediction of DOY of xylogenesis; for the remaining models we converted the predicted forcing, chilling and/or photoperiod at xylogenesis into a DOY by selecting the first DOY where the Forcing (or Chilling) thresholds were exceeded (or a value of 303 when chill thresholds not exceeded/365 when forcing sums never exceeded). Since daily photoperiods repeat twice per year (during the spring and fall seasons), for the photoperiod models we selected for each measurement -- the predicted DOY which was closer to the observed DOY (in most cases this was the spring prediction). The output of these models could be used directly in earth system models, in contrast to the regression models of Huang et al which follow the general form ($DOY_{wf} = \alpha + \beta * Photoperiod_{doy}$), and thus do not yield a unique solution to forecast DOY_{wf} in a given year/location but rather 365 predictions

All analyses were performed in R using the following packages: colorRamps, data.table, dplyr, egg, ggplot2, grid, insol, lme4, lubridate, Metrics,RColorBrewer, rlang, scales,tidyverse.

Table S1. Summary of all linear mixed models used to predict DOY_{wf} based on block cross validation of Huang et al's data. The first 5 models are illustrated in Figure 2; the remainder in Figure S2. For ease of comparison we have also included columns here which indicates the proportion of the density plots in each panel that is to the right of the 0 line in the corresponding figures.

Model	y	fixed effects	random effects	model prediction	Predicted DOY_{wf}	Proportion out performed by NULL model (spatial)	Proportion outperformed by NULL model (temporal)
forcing	FU_{wf}	Intercept	site, species	$FU_{wf(site,\ species,\ year)}$	min (doy) where $FU_{doy(site,\ year)} > FU_{wf(site,\ species,\ year)}$)	0.5	0.98
chilling	CU_{wf}	Intercept	site, species	$CU_{wf(site,\ species,\ year)}$	min (doy) where $CU_{doy(site,\ year)} > CU_{wf(site,\ species,\ year)}$)	1	1
photoperiod	$Photoperiod_{wf}$	Intercept	site, species	$Photoperiod_{wf(site,\ species,\ year)}$	min (doy) where $Photoperiod_{doy(site,\ year)} > Photoperiod_{wf(site,\ species,\ year)}$ & doy <173) OR min (doy) where $Photoperiod_{doy(site,\ year)} < Photoperiod_{wf(site,\ species,\ year)}$ & doy >173)	0.9	0.75
latitude	DOY_{wf}	latitude	site, species	$DOY_{wf(site,\ species)}$	$DOY_{wf(site,\ species)}$	0.66	0.86
MAT	DOY_{wf}	MAT	site, species	$DOY_{wf(site,\ species)}$	$DOY_{wf(site,\ species)}$	0.04	0.9
forcing_by_lat	FU_{wf}	latitude	site, species	$FU_{wf(site,\ species,\ year)}$	min (doy) where $FU_{doy(site,\ year)} > FU_{wf(site,\ species,\ year)}$)	0.49	0.98
forcing_by_chill	FU_{wf}	$CU_{April\ 30}$	site, species	$FU_{wf(site,\ species,\ year)}$	min (doy) where $FU_{doy(site,\ year)} > FU_{wf(site,\ species,\ year)}$)	0.39	0.99
forcing_by_chill_and_latitude	FU_{wf}	latitude +$CU_{April\ 30}$	site, species	$FU_{wf(site,\ species,\ year)}$	min (doy) where $FU_{doy(site,\ year)} > FU_{wf(site,\ species,\ year)}$)	0.43	0.99
photoperiod_by_MAT	$Photoperiod_{wf}$	MAT	site, species	$Photoperiod_{wf(site,\ species,\ year)}$	min (doy) where $Photoperiod_{doy(site,\ year)} > Photoperiod_{wf(site,\ species,\ year)}$ & doy <173) OR min (doy) where $Photoperiod_{doy(site,\ year)} < Photoperiod_{wf(site,\ species,\ year)}$ & doy >173)	0.3	0.77
photoperiod_by_lat	$Photoperiod_{wf}$	latitude	site, species	$Photoperiod_{wf(site,\ species,\ year)}$	min (doy) where $Photoperiod_{doy(site,\ year)} > Photoperiod_{wf(site,\ species,\ year)}$ & doy <173) OR min (doy) where $Photoperiod_{doy(site,\ year)} < Photoperiod_{wf(site,\ species,\ year)}$ & doy >173)	0.75	0.69
null	DOY_{wf}	Intercept	site, species	$DOY_{wf(site,\ species)}$	$DOY_{wf(site,\ species)}$		

Figure S2. Predictive performance of DOY_{wf} based on additional models of forcing and photoperiod (see Table S1). For each type of model, densities show the distribution of pairwise differences in RMSE prediction error for each train/evaluation subset vs predictions from the null model on the same subset ($\delta_{rmse(model\text{-}null)}$); a difference of 0 indicates the environmental drivers explain no additional variation beyond that of site and species identity; negative $\delta_{rmse(model\text{-}null)}$ values indicate better predictive performance than the null model and positive worse.



a
spatial
b
temporal
density
forcing by chill
forcing by chill and lat
forcing by lat
photoperiod by lat
photoperiod by MAT
0.12
0.09
0.06
0.03
0.00
0.4
0.2
0.0
-10
0
10
2
4
δrmse(model–null)
better ↔ worse

SI References

1. J.-G. Huang *et al.,* Photoperiod and temperature as dominant environmental drivers triggering secondary growth resumption in Northern Hemisphere conifers. Proc. Natl. Acad. Sci. U.S.A. *117*, 20645–20652 (2020).

2. D. Basler, Evaluating phenological models for the prediction of leaf-out dates in six temperate tree species across central Europe. *Agricultural and Forest Meteorology*, *217*, 10-21 (2016).

3. T.M. Crimmins, M.A. Crimmins, K.L Gerst, A.H. Rosemartin, J.F. Weltzin, USA National Phenology Network's volunteer-contributed observations yield predictive models of phenological transitions. *PLOS ONE* 12(8): e0182919 (2017).

4. Copernicus Climate Change Service (C3S) (2017): ERA5: Fifth generation of ECMWF atmospheric reanalyses of the global climate. Copernicus Climate Change Service Climate Data Store (CDS), (accessed 2020-08-20), https://cds.climate.copernicus.eu/cdsapp#!/home