* Discuss: Consistent CIs to use.

Regional Risk: Supplement

* ← more FS → less FS on figures?

Authors:

C. J. Chamberlain [1,2], B. I. Cook [3], I. Morales Castilla [1,4] & E. M. Wolkovich [1,2] 3← Add UBC please?

Author affiliations:

[1]Arnold Arboretum of Harvard University, 1300 Centre Street, Boston, Massachusetts, USA;

[2]Organismic & Evolutionary Biology, Harvard University, 26 Oxford Street, Cambridge, Massachusetts, USA;

[3]NASA Goddard Institute for Space Studies, New York, New York, USA;

[4]Edificio Ciencias, Campus Universitario 28805 AlcalÃą de Henares, Madrid, Spain

*Corresponding author: 248.953.0189; cchamberlain@g.harvard.edu

Methods: Spatial predictor

Spatial autocorrelation (SA) is a common issue in spatial ecology given that ~~close~~ nearby spatial units tend to be more similar than units far apart, and thus, cannot be considered as independent units, which is a frequent assumption in statistical tests (Diniz-Filho *et al.*, 2003). If model residuals are spatially autocorrelated, and thus, non-independent then model coefficients and errors may be biased in a hard-to-predict way (Mauricio Bini *et al.*, 2009). On the contrary, if model residuals are notautocorrelated, then SA should not be of concern (Hawkins, 2012).

To control for spatial autocorrelation and to account for spatially structured processes independent from our regional predictors of false springs, we generate an additional *spatial predictor* for the model. To avoid collinearity, we computed our *spatial predictor* from the residuals of a linear model of false springs as a function of all other regional factors that are also spatially structured (e.g. spring temperature, altitude, distance to the coast), following the logic of spatial filter modelling (Diniz-Filho & Bini, 2005). The calculation of the *spatial predictor* followed the next steps: (a) we fit a linear model of false spring versus regional factors, Fix

$$
\begin{aligned}
y_i \sim N(\alpha(i)) &+ \beta_{NAO_{(i)}} + \beta_{MeanSpringTemp_{(i)}} + \beta_{Elevation_{(i)}} + \beta_{DistanceCoast_{(i)}} \\
&+ \beta_{ClimateChange_{(i)}} + \beta_{NAO \times Species_{(i)}} + \beta_{MeanSpringTemp \times Species_{(i)}} + \beta_{Elevation \times Species_{(i)}} \\
&+ \beta_{DistanceCoast \times Species_{(i)}} + \beta_{ClimateChange \times Species_{(i)}} \\
&+ \beta_{NAO \times ClimateChange_{(i)}} + \beta_{MeanSpringTemp \times ClimateChange_{(i)}} + \beta_{Elevation \times ClimateChange_{(i)}} \\
&+ \beta_{DistanceCoast \times ClimateChange_{(i)}} + \sigma_{sp_{(i)}}
\end{aligned}
\tag{S1}
$$

?.?

eqn should not break up text

(b) We extracted the residuals of the regression Equation S1, which represent the portion of the variation in the number of false springs that is independent from the predictors in the model. (c) Residuals were utilized as our Y values in a selection of spatial eigenvectors aimed at keeping only the minimal subset of spatial eigenvectors that are able to remove SA from model residuals. Specifically, we selected eigenvectors following the the minimization of Moran's I of the residuals (MIR) approach (Griffith & Peres-Neto, 2006; Diniz-Filho *et al.*, 2012; David *et al.*, 2017). (d) We fit a linear model between the residuals of Equation S1 and the subset of selected eigenvectors. And (e) we ~~take~~ took the fitted values from this regression as our *spatial predictor* in our final model (see equation from main text), which can be interpreted as a latent variable summarizing the spatial structure in false springs that is unaccounted for by the rest of regional factors in our model (Morales-Castilla *et al.*, 2012). A *spatial predictor* generated in this way has three major advantages. First, it ensures that no SA is left in model residuals. Second, it avoids introducing collinearity issues with other predictors in the model. And third, it can be interpreted as a latent variable summarizing spatial processes (e.g. local adaptation, plasticity, etc.) occurring at multiple scales.

methods are past tense

We wanted to X, but Because there is no data of type y...

Species rate of budburst calculations

We used data from a growth chamber experiment (Flynn & Wolkovich, 2018) to determine the average number of days between budburst and leafout for our study species. Cuttings for the experiment were made in January 2015 from two field sites: Harvard Forest (HF, 42.5°N, 72.2°W) and the Station de Biologie des Laurentides in St-Hippolyte, Québec (SH, 45.9°N, 74.0°W). The experiment examined budburst and leafout for *Acer saccharum* (Marshall), *Alnus incana* (L.), *Betula papyrifera* (Marshall), *Fagus grandifolia* (Ehrh.), *Fraxinus nigra* (Marshall), and *Quercus alba* (L.) in a fully crossed design of three levels of chilling (field chilling, field chilling plus 30 days at either 1 or 4 °C), two levels of forcing (20°C/10°C or 15°C/5°C day/night temperatures, such that thermoperiodicity followed photoperiod) and two levels of photoperiod (8 versus 12 hour days) resulting in 12 treatment combinations. Phenological observations of each cutting were made every 2-3 days over 82 days. Phenology was assessed using a BBCH scale that was modified for trees

think you can just ref paper

(Finn *et al.*, 2007). We then took the mean number of days between budburst and leafout for the entire experiment, which was 12 days. We compared this number to a field observation study (Donnelly *et al.*, 2017) that looked at the time between budburst and leafout across 10 species over 5 years. Finally, data were provided by the USA National Phenology Network and the many participants who contribute to its NatureâĂŹs Notebook program (USA-NPN,2019; www.usanpn.org/data/observational) for *Aesculus flava* (Sol.), *Aesculus glabra* (Willd.), *Alnus incana* (Moench.), *Betula nigra* (L.), *Betula papyrifera* (Marshall), *Fagus grandifolia* (Ehrh.), *Fraxinus americana* (L.), *Fraxinus nigra* (Marshall) and *Quercus velutina* (Lam.) and took the mean number of days between budburst and leafout. Across all three approaches, the average duration of vegetative risk was approximately 12 days.

X

you're switching terminology too much for regular reader

To determine varying durations of vegetative risk for each species we used data from the growth chamber experiment (Flynn & Wolkovich, 2018). We used the rate of budburst of *Acer saccharum* (Marshall) for *Aesculus hippocastanum* (Buerki *et al.*, 2010), *Alnus incana* for *Alnus glutinosa*, *Betula papyrifera* for *Betula pendula* (Wang *et al.*, 2016), *Fagus grandifolia* for *Fagus sylvatica*, *Fraxinus nigra* for *Fraxinus excelsior* and *Quercus alba* (L.) for *Quercus robur* (Hipp *et al.*, 2017).

? how diff. from above?

Results: The effects of climatic and spatial variation on false spring incidence

The overall model output estimates are for *Aesculus hippocastanum* as species were used as two-way interactions to simulate modeled groups on the main effects. The model estimates on the logit scale were converted to probability percentages for easier interpretation. To convert we used the equations below (Gelman & Hill, 2006):

$$inverselogit = (1/(1+exp(-(ModelEstimate)))) \tag{S2}$$

$$\begin{aligned} &inverselogit(Intercept + (ModelEstimate/(sd(RawDataPredictor)*2))*mean(RawDataPredictor)) - \\ &inverselogit(Intercept + (ModelEstimate/(sd(RawDataPredictor)*2)) \\ &*(mean(RawDataPredictor) - (2*sd(RawDataPredictor)))*100 \end{aligned} \tag{S3}$$

not tex math

Good!

FS?

Supplement: Tables and Figures

Table 1: Data collected from PEP725 for each species

Species	Num. of Observations	Num. of Sites	Num. of Years
Aesculus hippocastanum	156468	10157	66
Alnus glutinosa	91094	6775	65
Betula pendula	154897	10139	66
Fagus sylvatica	129133	9099	66
Fraxinus excelsior	92665	7327	65
Quercus robur	131635	8811	66

from model?

credible?

Table 2: Mean budburst days and ~~confidence~~ intervals for each species for before (1951-1983) and after climate change (1984-2016).

1980s+ 1980s+

change formatting?

Species (years)	Mean Budburst	2.5%	97.5%
Aesculus hippocastanum (1984-2016)	95.35	95.26	95.44
Alnus glutinosa (1984-2016)	94.90	94.67	95.13
Betula pendula (1984-2016)	95.44	95.23	95.66
Fagus sylvatica (1984-2016)	103.75	103.52	103.97
Fraxinus excelsior (1984-2016)	113.48	113.26	113.71
Quercus robur (1984-2016)	109.60	109.38	109.82
Aesculus hippocastanum (1951-1983)	102.20	102.00	102.41
Alnus glutinosa (1951-1983)	102.81	102.27	103.36
Betula pendula (1951-1983)	101.31	100.81	101.81
Fagus sylvatica (1951-1983)	109.07	108.56	109.59
Fraxinus excelsior (1951-1983)	119.36	118.82	119.89
Quercus robur (1951-1983)	115.85	115.34	116.36

always give 50% + 2.5% + 97.5%

... otherwise your readers focus on the WRONG thing.

super unlikely?

Table 3: Summary of Bernouilli model of false spring risk without the species interactions (estimates presented on logit scale for *Aesculus hippocastanum*).

Term	Model Estimate	10%	90%
NAO Index	0.14	~~0.12~~	0.16
Mean Spring Temperature	-0.48	-0.50	-0.45
Distance from Coast	0.40	0.38	0.43
Elevation	0.19	0.17	0.22
Space Parameter	-0.06	-0.08	-0.04
Climate Change	0.35	0.33	0.37
NAO Index by Climate Change	-0.83	-0.85	-0.81
Mean Spring Temperature by Climate Change	0.42	0.40	0.44
Distance from Coast by Climate Change	-0.12	-0.15	-0.10
Elevation by Climate Change	0.00	-0.03	0.03
Space Parameter by Climate Change	-0.05	-0.07	-0.03

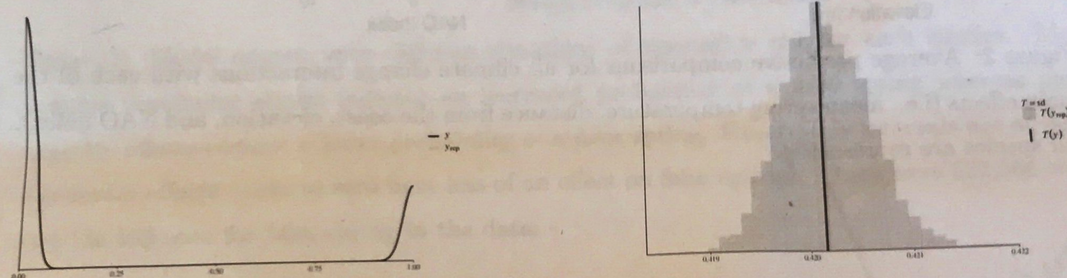
be consistent!

50%

$\boxed{25\% \quad 75\%}$

2.5%

97.5%



(a) (b)

Figure 1: (a) Posterior predictive check comparing the simulated model estimates to the raw data. The curves overlap greatly, which suggests our model is valid and fit the data. (b) Posterior predictive check comparing the standard deviation from our model output to the data. The model fits our data well, which suggests our model is valid.

not needed, cut.

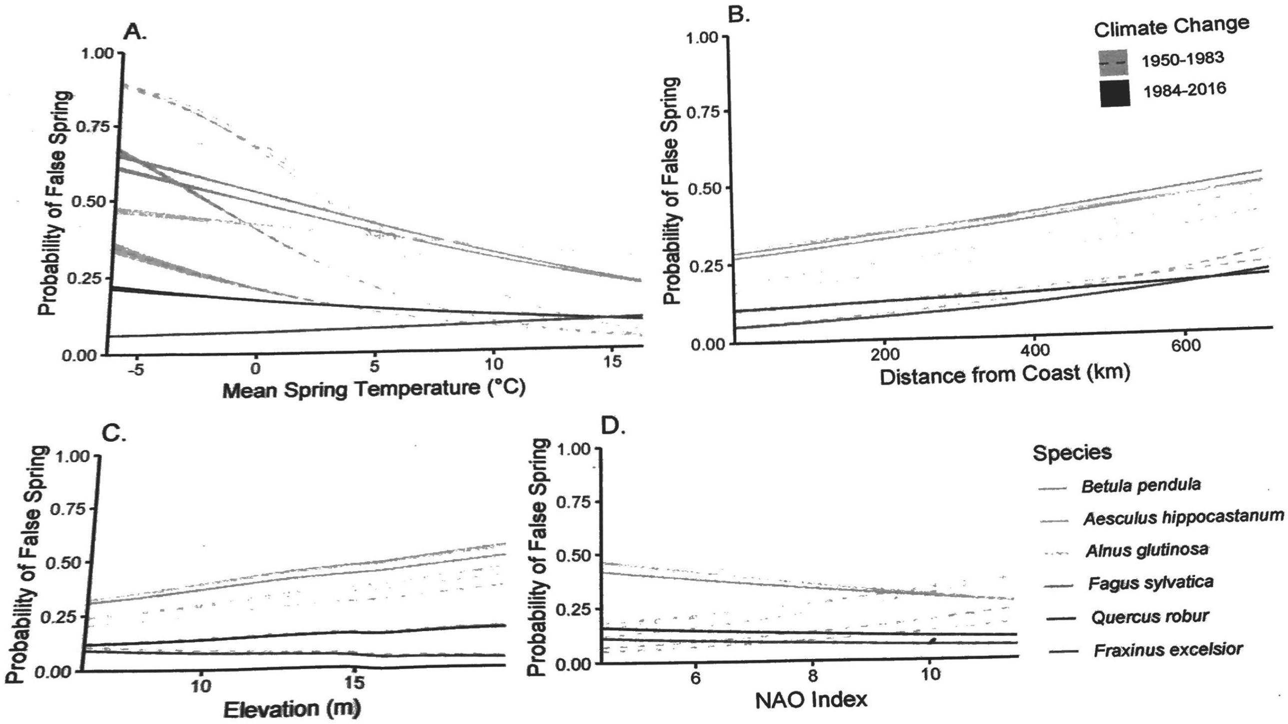


Figure 2: **Average predictive comparisons for all climate change interactions with each of the main effects (i.e., mean spring temperature, distance from the coast, elevation, and NAO index). All species are represented.**

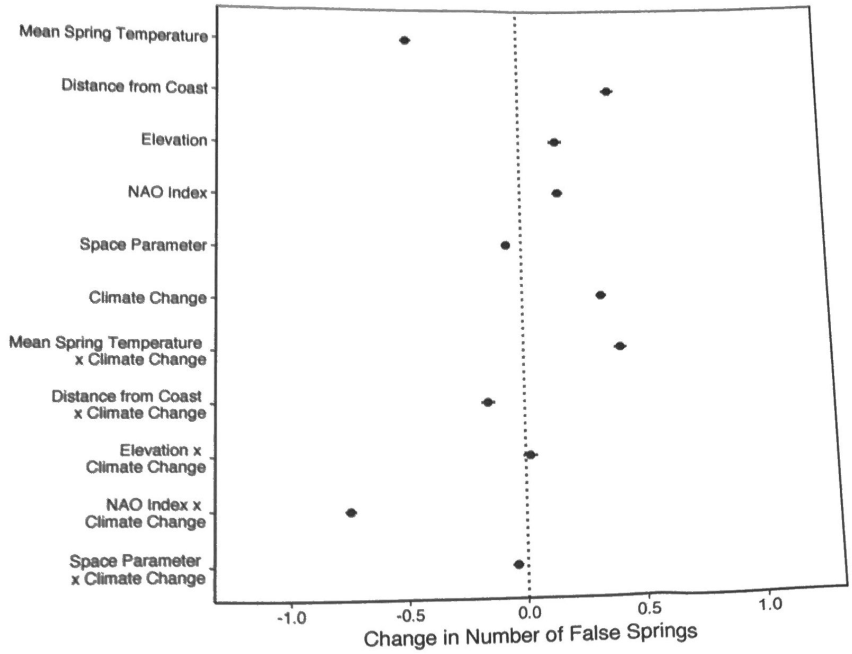


Figure 3: Model output with different durations of vegetative risk for each species. More positive parameter effects indicate an increased probability of a false spring whereas more negative effects suggest a lower probability of a false spring. Uncertainly intervals are at 90%. →10%? Parameter effects closer to zero have less of an effect on false springs. There were 622,565 zeros and 132,463 ones for false spring in the data.

see comment next pg.

uncertainty/credible = either is fine but be consistent

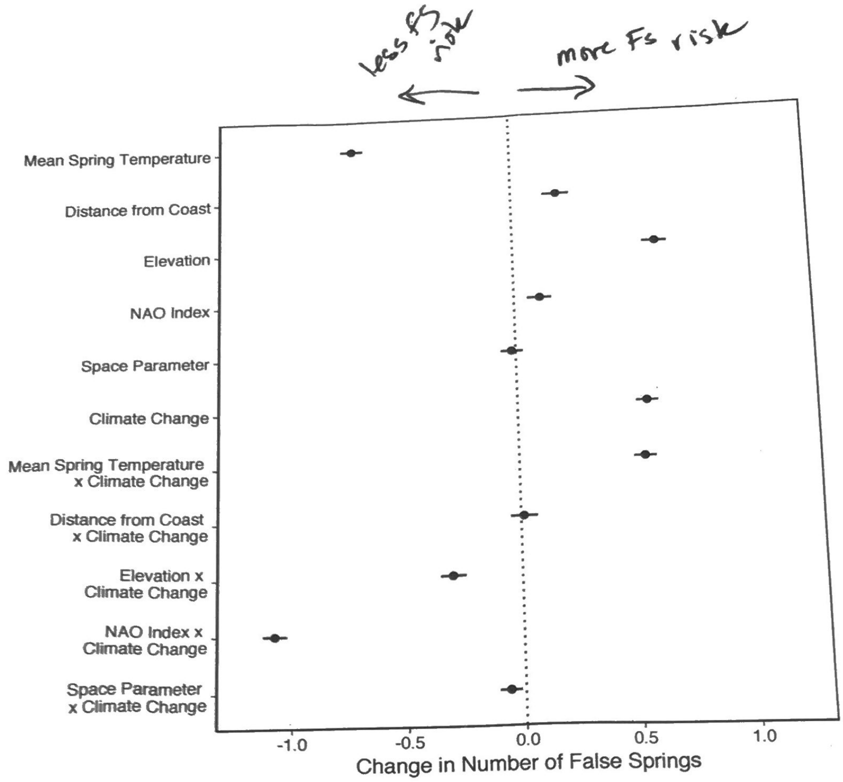


Figure 4: Model output with a lower temperature threshold (-5°C) for defining a false spring. More positive parameter effects indicate an increased probability of a false spring whereas more negative effects suggest a lower probability of a false spring. Uncertainly intervals are at 90%. Parameter effects closer to zero have less of an effect on false springs. There were 730,996 zeros and 23,855 ones for false spring in the data, rendering a less stable model.

x10%

?

Should we just add arrows to this + all similar figures? Would be nicer?