Supplemental Materials for: How do climate change experiments alter local climate?

A.K. Ettinger, I. Chuine, B. Cook, J. Dukes, A.M. Ellison, M.R. Johnston, A.M. Panetta, C. Rollinson, Y. Vitasse, E. Wolkovich

April 26, 2018

Additional methods for database development

For our literature review, we searched Web of Science (ISI) for Topic=(warm* OR temperature*) AND Topic=(plant* AND phenolog*) AND Topic=(experiment* OR manip*). We restricted dates to the time period after the STONE database (i.e. January 2011 through March 2015). This yielded 277 new studies. We removed all passive warming studies from the list, and contacted authors for daily data. Three additional datasets were offered or suggested to us, and in March 2018, we added additional studies found by using the same terms to search the following online databases for additional datasets: dataONE (dat, 2018), KNB (knb, 2018), and dryad (dry, 2018). The resulting database contains daily climate data collected between 1991 and 2015 from 15 North American, European, and Chinese climate change experiments (Table S1, Figure S1).

Details of statistical analyses and results

For all analyses, we use mixed-effects models implemented using the lme4 package in R, version 3.2.4 (Bates et al., 2015; Team, 2016). Mixed-effects models, also called multi-level or hierarchical models, can account for structured data that violate the independence assumption of traditional linear regression (Gelman and Hill, 2007). In our analyses, we use levels/groupings of experimental site, year, and day of year (doy) to account for this mutual dependence among data points. To test for significance of fixed effects in our models, we use Type II tests for models including only main effects and Type III tests for models including interactions, as well as main effects.

Analysis of effects of warming on daily temperature range

To test how active warming alters daily temperature range (DTR, the difference between maximum and minimum temperatures in a day), we used data from the 10 studies in the MC3E database that include daily measurements of soil and/or above-ground (i.e., air, canopy, surface) temperature maxima and minima. For consistency, we included only structural controls (we therefore excluded exp15, which used only ambient controls), and also excluded data from plots with precipitation treatments. We fit linear mixed-effect models with above-ground DTR, soil DTR, minimum and maximum daily above-ground temperature, and minimum and maximum daily soil temperature as response predictors. We included target temperature treatment (or measured temperature, for those studies that did not have explicit target temepratures) as a fixed effect. Random effects were site and study year nested within site (with random slopes and intercepts, Tables S3 & S4).

Analysis of effects of time and space on experimental microclimate

To test how treatment effects vary spatially (i.e., among blocks within a study) and temporally (i.e., among years within a study), we used data from the four studies in the MC3E database that used blocked designs. We fit linear mixed-effect models with mean daily soil temperature, minimum daily air temperature, and maximum daily air temperature as response predictors (Figure 3 in the main text). For temporal models, we included fixed effects of temperature treatment, year, and their interaction; random effects were site and block nested within site (intercept-only structure, Table **??**). For spatial models, we included fixed effects of temperature treatment, block, and their interaction; random effects were site and year nested within site (intercept-only structure, Table **??**). Both of these models excluded data from plots with precipitation treatments.

Analysis of effects of infrastructure on experimental microclimate

To test how infrastructure affects local climate, we compared temperature and soil moisture data from the studies in the C3E database that monitored climate in two types of control plots: structural controls (i.e., 'shams' or 'disturbance controls,'which contained all the warming infrastructure, such as soil cables or forced air chambers with no heat applied) and ambient controls with no infrastructure added. These five studies consisted of soil and forced air warming types and occurred at two sites: Duke Forest and Harvard Forest (Farnsworth et al., 1995; Clark et al., 2014; Marchin et al., 2015; Pelini et al., 2011). We fit linear mixed effects models by month for the following response variables: mean daily soil temperature, minimum and maximum daily air and soil temperature (Farnsworth et al. (1995) only measured mean soil, not minimum and maximum air or soil so there are only four different studies in those models), and soil moisture. The predictor was control type (sham or ambient). To allow for both mean differences in temperature and the effect of control to to vary among sites and years, random effects were site and year nested within site, modeled with a random slopes and random intercepts structure. We found that experimental structures altered above-ground and soil temperatures in opposing ways: above-ground temperatures were higher in the structural controls, compared with ambient conditions with no structures installed, whereas soil temperatures were lower in the structural controls compared with ambient soil (Figure 4 in the main text). In addition, soil moisture was lower in structural controls compared with ambient conditions. These general patterns were consistent across the different temperature models we fit (mean, minimum, and maximum soil and air temperatures), although the magnitude varied across months, as well as among studies. We show summaries from models fit to the entire year (Tables **??**, **??**, **??**), as well as summaries from models fit to each month of data, as is shown in Figure 4 in the main text (Tables **??**, **??**, **??**).

No infrared heating studies in our database included both ambient and structural controls, so we were unable include this warming type in the above analysis. We conducted a separate analysis of infrared studies, taking advantage of the fact that some studies in the database used ambient controls and some used structural controls. We focused on the seven infrared studies that measured soil temperature; three of these experiments used ambient controls (exp06, exp11, exp14) and four used structural controls (exp1,exp09,exp12, exp13). To understand whether infrared heating equipment is likely to alter microclimate, we calculated Tdiff, the difference between mean annual temperature in control plots and in each treatment level within a study. We then divided by target warming of the treatment, to standardize across studies. We expected that, if control type affects microclimate, there should be strong differences across the two control types in the amount of warming achieved per degree of target warming. We tested for these differences betwe fitting a mixed effects model with Tdiff as the response variable, control type as the explanatory variable, and random effects of site and year nested within site (Table **??**). Though perhaps with only weak inference, this allowed us to explore the potential effects of heating equipment on microclimate in infrared studies.

Analysis of effects of precipitation treatments on above-ground temperature

Of the 15 experiments in the C3E database, four manipulated precipitation and measured above-ground temperature and three of these also measured soil temperature. To examine the effects of precipitation treatment on temperature, we fit linear mixed effect models to data from these sites with temperature (above-ground daily minimum and maximum, and soil minimum and maximum) as the response variables. Predictors were precipitation treatment (a continuous fixed effect, which ranged from 50 to 200 % of ambient for these four studies), target warming (a continuous fixed effect, which ranged from 0 to 4 °C for thesefour studies), and their interaction. To account for methodological and other differences among site, we included site as a random effect, with year and doy nested within site to account for the non-independent nature of measurements taken on the same day within sites. We used a random intercept model structure, (Table **??**).

Analysis of effects of experimental warming on soil moisture

Of the 15 experiments in the C3E database, 12 measured and reported soil moisture. To examine the effects of target warming treatment on soil moisture, we fit linear mixed effects models to data from these ten sites, excluding plots with precipitation treatments. We first fit a model with soil moisture as the response and a predictor of target warming (this was a continuous fixed effect, which ranged from 0 to 5.5 °C for these 12 studies). To account for methodological and other differences among site, we included site as a random effect, with year and doy nested within site to account for the non-independent nature of measurements taken on the same day within sites. We used a random slope and intercept model structure, to allow the effect of target warming to vary among sites (Table **??**).

In addition to testing how experimental warming influenced soil moisture, we also tested how experimental structures influenced soil moisture. We compared the soil moisture measured in structural controls to both ambient controls and warmed plots by fitting a model with categorical fixed effects of “ambient,” “structural control,” and “warmed.” We again included site as a random effect, with doy nested within site to account for the non-independent nature of measurements taken on the same day within sites, and used a random intercept structure (Table **??**).

Analysis of budburst phenology

We wanted to investigate how using measured plot-level climate variables, as opposed to target warming, alters estimates of temperature sensitivity in ecology. To do this, we fit two different types of models to data from the five study sites in the C3E database that recorded above-ground temperature and soil moisture, as well as phenology data (doy of budburst). We focus on budburst, as this phenological phase was reported most commonly among studies in the C3E database. For all models, we accounted for non-independence by including species, site, and year nested within site as intercept-only random effects (Table **??**). The target warming model included only one explanatory variables (the target amount of warming). We compared this to models with mean annual measured above-ground temperature (offset by subtracting the minimum temperature across all studies and plots, to make model intercepts more similar), mean winter (January-March) soil moisture, and their interaction as explanatory variables. The slope for temperature in the measured climate model can be directly compared to the slope for target warming in the target warming model because the units are the same (change in budburst, in days/°C).

To determine which specific above-ground temperature variable to include, we compared AICs of models for with four different temperature variables (mean annual minimum and maximum temperatures, mean January-March minimum and maximum temperatures). The model with mean annual minimum temperature, mean January-March soil moisture, and their interaction provided the best model fit (lowest AIC, highest explained variation, Table **??**), so we discuss and interpret that model in the main text, summarize it in Table **??**, and present its coefficients in Figure 7.

References

2018. Data observation network for earth (dataone).

2018. Dryad digital repository.

2018. Knowledge network for biocomplexity (knb).

Bates, D., M. Maechler, B. M. Bolker, and S. Walker. 2015. Fitting linear mixed-effects models using lme4. Journal of Statistical Software 67:1–48.

Clark, J. S., J. Melillo, J. Mohan, and C. Salk. 2014. The seasonal timing of warming that controls onset of the growing season. Global Change Biology 20:1136–1145.

Farnsworth, E., J. Nunez-Farfan, S. Careaga, and F. Bazzaz. 1995. Phenology and growth of three temperate forest life forms in response to artificial soil warming. Journal of Ecology 83:967–977.

Gelman, A., and J. Hill. 2007. Data Analysis Using Regression and Multilevel/Hierarchical Models. Cambridge University Press, New York, NY, USA.

Marchin, R. M., C. F. Salk, W. A. Hoffmann, and R. R. Dunn. 2015. Temperature alone does not explain phenological variation of diverse temperate plants under experimental warming. Global Change Biology 21:3138–3151.

Pelini, S. L., F. P. Bowles, A. M. Ellison, N. J. Gotelli, N. J. Sanders, and R. R. Dunn. 2011. Heating up the forest: open-top chamber warming manipulation of arthropod communities at Harvard and Duke Forests. Methods in Ecology and Evolution 2:534–540.

Team, R. C. 2016. R: A language and environment for statistical computing.

Supplemental Tables

Table S1: **Experimental sites included in the C3E database**. Experimental sites correspond to the map (Figure 1, main text). We give the study ID, location, source, years of data included, warming type, the type of control used to maintain warming, plot area (in m2), watts of heating output, target warming treatment (°C), precipitation treatment (proportion of ambient), method of above-ground temperature measurement (with height of measurement, in cm, for air), depth of soil temperature measurement (cm), depth of soil moisture measurement (cm) used in each study, and statistical analysis used in the source listed (i.e., ANOVA with categorical explatory variables for different warming levels versus continuous microclimate explanatory variables). Note that some sites may have multiple sources; however, we list only one here.

study	location	data years	source	warming type	warming control	area	watts	warming trtmt	precip trtmt	above-ground temp	soil temp depth	soil moist depth	control type	analysis type
exp01	Waltham, MA, USA	Hoeppner and Dukes 2012	2009-2011	infrared	constant, feedback	4.00	50, 150, 250	1, 2.7, 4	0.5, 1.0, 1.5	canopy	2, 10	30	structural	categorical
exp02	Montpelier, France	Morin et al. 2010	2004	infrared	feedback	1.56	102.4	1.5, 3	0.7, 1.0			15, 30	ambient	categorical
exp03	Duke Forest, NC, USA	Clark et al. 2014	2009-2014	forced air and soil	feedback	17.00		3, 5		air (30)	10	30	both	continuous
exp04	Harvard Forest, MA, USA	Clark et al. 2014	2009-2012	forced air and soil	feedback	17.00		3, 5		air (30)	10	30	both	continuous
exp05	Jasper Ridge Biological Preserve, CA, USA	Cleland et al. 2007	1998-2002	infrared	constant	3.14	80	1.5	1.0, 1.5		15	15	ambient	categorical
exp06	Rocky Mountain Biological Lab, CO, USA	Dunne et al. 2003	1995-1998	infrared		30.00	22	1.9			12, 25	12, 25	ambient	categorical
exp07	Harvard Forest, MA, USA	Pelini et al. 2011	2010-2015	forced air	feedback	15.70		1.5-5.5		air (22)	2, 6	30	both	continuous
exp08	Harvard Forest, MA, USA	Farnsworth et al. 1995	1993	soil warming		36.00		5			5		both	categorical
exp09	Stone Valley Forest, PA, USA	Rollinson and Kaye 2012	2009-2010	infrared	feedback	4.00	100	2	1.0, 1.2	surface	3	8	structural	categorical
exp10	Duke Forest, NC, USA	Marchin et al. 2015	2010-2013	forced air	feedback	15.70		1.5-5.5		air (22)	2, 6	30	both	continuous
exp11	Rocky Mountain Biological Lab, CO, USA	Price and Wasser 1998	1991-1994	infrared	constant	30.00	15	1.2			12		ambient	categorical
exp12	Kessler Farm Field Laboratory, OK, USA	Sherry et al. 2007	2003	infrared	constant	6.00	100	4.17	1.0, 2.0	air (14)	7.5, 22.5	15	structural	categorical
exp13	Haibei Alpine Grassland Research Station, China	Suonan et al. 2017	2012-2014	infrared	constant	3.96	303	1.5		air (30)	5, 10	5, 10	structural	categorical
exp14	Cedar Creek, MN, USA	Whittington et al 2015	2009-2011	infrared	constant	7.50	80, 133	1.5, 3		air (10,25)	1, 10	6	ambient	categorical
exp15	Oak Ridge, TN, USA	Gunderson et al 2015	2003-2005	forced air	feedback	9.42		2, 4		air		10, 20	both	categorical

Table S2: **Summary of warming treatments, by warming type** for studies included in the MC3E database, and for another common type of experimental warming, passive open-top chambers (OTC), for comparison. Summaries of the target warming treatments (°C) and measured warming for air temperature, soil surface temperature, and soil temperture are given. Measured warming shown here is for warming treatments only (precipitation treatments are expcluded), and is the difference between mean annual temperature (MAT) of control plots and MAT of each treatment level within a study. Mean difference (with standard error) and the range of differences (minimum to maximum differences in MAT) is shown. n is the number of studies of each warming type in the MC3E database.

	otc	force_air	force_air_soil	infrared	soil
target (min-max)	NA	3.5 (1.5-5.5)	4 (3-5)	2.2 (1-4)	5 (5-5)
air_mean (se)	0.8 (0.1)	3.22 (0.12)	1.9 (0.24)	1.08 (0.16)	
air_range	0.5-1.3	0.83-5.2	-0.05-3.59	0.42-1.83	
surf_mean (se)	0.9 (0.1)			1.72 (0.1)	
surf_range	0.4-1.4			1.52-1.87	
soil_mean (se)	0.8 (0.3)	1.29 (0.09)	3.01 (0.34)	1.73 (0.2)	5.04 (NA)
soil_range	-0.1-3.9	-0.93-3.46	0.02-5.08	-0.06-4.31	NA
n	0 (from Bokhorst et al. 2013)	2	2	9	1

Table S3: **Summary of linear mixed-effects models of how target warming treatment affects daily temperature range, mimumim, and maximum above-ground temperatures in climate change experiments.** We excluded data from plots with precipitation treatments from these analyses. Estimates (est.) are the intercept and coefficient for target warming from the model; se is the standard error for these estimates. The effect of target warming on observed warming was significant based on Type II Wald χ^2 tests of fixed effects for minimum above-ground temperature (χ^2=46.58, df=1, p<0.001) and maximum above-ground temperature (χ^2=5.74, df=1, p=0.02), but not for daily temperature range (DTR) (χ^2=1.38, df=1, p=0.24). Random effects were site (n=8) and year nested within site (n=32 year-site combinations), with a random slope and intercept structure, and are listed for each model. Total number of observations=169,797, and units are °C for all three models.

	above-ground DTR		above-ground min temp.		above-ground max temp.	
predictor	est.	se	est.	se	est.	se
intercept	14.74	1.58	6.5	0.89	21.46	1.97
target warming effect	-0.37	0.31	0.81	0.12	0.48	0.2
site random effects	int	target	int	target	int	target
exp01	13.65	0.06	6.48	0.8	19.94	0.85
exp03	12.65	-0.13	7.99	0.49	20.95	0.38
exp04	9.5	0.01	6.36	0.49	15.68	0.49
exp07	10.28	0.37	4.61	0.72	14.04	1.06
exp09	17.37	-0.61	4.72	1.13	21.28	0.56
exp10	12	0.23	8.58	0.73	21.44	0.97
exp12	22.63	-2.57	4.85	1.43	29.65	-0.82
exp14	19.84	-0.3	8.39	0.68	28.68	0.37

Table S4: **Summary of linear mixed-effects models of how target warming treatment affects daily temperature range, mimumim, and maximum above-ground temperatures in climate change experiments.** We excluded data from plots with precipitation treatments from these analyses. Estimates (est.) are the intercept and coefficient for target warming from the model; se is the standard error for these estimates. The effect of target warming on observed warming was significant based on Type II Wald χ^2 tests of fixed effects for minimum above-ground temperature (χ^2=46.58, df=1, p<0.001) and maximum above-ground temperature (χ^2=5.74, df=1, p=0.02), but not for daily temperature range (DTR) (χ^2=1.38, df=1, p=0.24). Random effects were site (n=8) and year nested within site (n=32 year-site combinations), with a random slope and intercept structure, and are listed for each model. Total number of observations=169,797, and units are °C for all three models.

	soil DTR		soil min temp.		soil max temp.	
predictor	est.	se	est.	se	est.	se
(Intercept)	4.13	0.47	10.38	1.07	14.44	1.37
target	0.02	0.07	0.72	0.1	0.75	0.09
site random effects	int	target	int	target	int	target
exp01	6.45	0.24	10.29	0.79	16.29	0.96
exp03	2.83	-0.01	12.35	0.74	15.24	0.74
exp04	2.13	0.06	9.17	0.87	11.12	0.92
exp06	4.19	0.02	7.97	0.68	12.02	0.76
exp07	2.72	0.05	7.28	0.41	10.27	0.51
exp09	3.62	-0.27	9.48	1.16	13.07	0.9
exp10	5.14	0.06	12.35	0.43	17.72	0.51
exp11	3.97	0.02	6.12	0.66	9.92	0.78
exp12	4.05	0.24	12.16	0.69	15.53	0.84
exp14	6.15	-0.22	16.66	0.76	23.17	0.55