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 5, 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 time and space on local experimental climate

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 C3E 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 S4). 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 S5). Both of these models excluded data from plots with precipitation treatments.

Analysis of effects of infrastructure on local experimental climate

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 S6, S7, S8), as well as summaries from models fit to each month of data, as is shown in Figure 4 in the main text (Tables S9, S10, S11).

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 S13).

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 S14).

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 S15).

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 S16). 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 S17), so we discuss and interpret that model in the main text, summarize it in Table S16, 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,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), and depth of soil moisture measurement (cm) used in each study. Note that some sites may have multiple sources; however, we list only one here.

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

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

Table S3: **Summary of linear mixed-effects models of how target warming treatment affects daily temperature range for above-ground temperatures, and for 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=40.95, df=1, p<0.001) and maximum above-ground temperature (χ^2=4.63, df=1, p=0.03), but not for daily temperature range (DTR) (χ^2=1.18, df=1, p=0.28). Random effects were site (n=7) and year nested within site (n=29 year-site combinations), with a random slope and intercept structure. Total number of observations=135,943, and units are °C for all three models.

	DTR		min above-ground temp.		max above-ground temp.	
predictor	est.	se	est.	se	est.	se
intercept	14.72	1.4	6.77	0.83	21.62	1.77
target warming effect	-0.35	0.28	0.84	0.11	0.53	0.18
site random effects	int	target	int	target	int	target
exp01	13.65	0.06	6.59	0.8	20.01	0.85
exp03	12.48	-0.08	8.05	0.49	20.81	0.41
exp04	9.44	0.02	6.19	0.52	15.48	0.53
exp07	10.57	0.3	4.35	0.77	14.12	1.05
exp09	17.37	-0.61	4.95	1.13	21.37	0.56
exp10	12.12	0.2	8.77	0.77	21.6	0.96
exp12	22.62	-2.56	5.56	1.44	29.6	-0.8
exp14	19.83	-0.3	8.63	0.7	28.83	0.37
exp15	14.37	-0.18	7.84	0.98	22.72	0.8

Table S4: **Analysis of variance table for temporal linear mixed-effects models of daily mean soil temperature, minimum above-ground temperature, and maximum above-ground temperature**, fit by maximum likelihood. See Figure 3 in the main text. We list degrees of freedom (which are identical across response variables), test statistics, and p-values for Type III Wald χ^2 tests of fixed effects in the models. For all models, random effects were site (n=4 for soil temperature model, n=3 for above-ground temperature models) and block nested within site (intercept-only structure; n=18 for soil, n=12 for above-ground); total number of observations=36,813 for soil and 28,875 for above-ground; units are °C.

		mean soil temp.		min above-ground temp.		max above-ground temp.	
predictor	df	χ^2	p	χ^2	p	χ^2	p
intercept	1	40.12	<0.001	42.93	<0.001	259.85	<0.001
temp. treatment	4	952.76	<0.001	369.2	<0.001	38.9	<0.001
year	5	506.78	<0.001	1044.39	<0.001	472.13	<0.001
temp. treatment:year	14	575.48	<0.001	314.94	<0.001	378.02	<0.001

Table S5: **Analysis of variance table for spatial linear mixed-effects models of daily mean soil temperature, minimum above-ground temperature, and maximum above-ground temperature**, fit by maximum likelihood. See Figure 3 in the main text. We list degrees of freedom (which are identical for all models), test statistics, and p-values for Type III Wald χ^2 tests of fixed effects in the models. For all models, random effects were site (n=6) and year nested within site (intercept-only structure; n=6); total number of observations=17,177.

		mean soil temp.		min above-ground temp.		max above-ground temp.	
predictor	df	χ^2	p	χ^2	p	χ^2	p
intercept	1	405.66	<0.001	61.45	<0.001	557.64	<0.001
temp. treatment	4	572.96	<0.001	317.11	<0.001	112.09	<0.001
block	39	405.84	<0.001	138.21	<0.001	171.78	<0.001
temp. treatment:block	80	468.61	<0.001	243.17	<0.001	388.57	<0.001

Table S6: **Summaries of linear mixed-effects models comparing effects of ambient versus structural controls on daily mean, minimum, and maximum soil temperature in climate change experiments across the year**. Estimates (est.) are the intercept (representing ambient controls) and coefficient (representing structure effects) from the models; se is the standard error for these estimates. For these annual models, differences between control types were significant based on Type II Wald χ^2 tests of fixed effects for mean soil temperature (χ^2=5.53, df=1, p=0.02) and minimum soil temperature (χ^2=3.87, df=1, p=0.05), but not for maximum soil temperature (χ^2=2.07, df=1, p=0.15). For all models, random effects of site (n=5 for mean model, n=4 for min and max models) and year nested within site (n=21 for mean model, n=20 for min and max models) were fit with a random slope and intercept structure; total number of observations= 48,860 for the mean model and 44,530 for the min and max models; units are °C.

	mean soil temp.		min soil temp.		max soil temp.	
predictor	est.	se	est.	se	est.	se
intercept	11.89	1.42	10.81	1.48	13.92	1.61
structure effect	-0.57	0.24	-0.63	0.32	-0.54	0.38

Table S7: **Summaries of linear mixed-effects models comparing effects of ambient versus structural controls on daily minimum and maximum air temperature in climate change experiments, across the year.** Estimates (est.) are the intercept (representing ambient controls) and coefficient (representing structure effects) from the models; se is the standard error for these estimates. For these annual models, differences between control types were not significant based on Type II Wald χ^2 tests of fixed effects for minimum air temperature (χ^2=1.07, df=1, p=0.30), nor for maximum air temperature (χ^2=0.01, df=1, p=0.91). For both models, random effects of site (n=4) and year nested within site (n=20) were fit with a random slope and intercept structure; total number of observations= 44,085; units are °C.

	min air temp.		max air temp.	
predictor	est.	se	est.	se
intercept	6.29	1.51	17.74	1.81
structure effect	0.36	0.35	0.02	0.21

Table S8: **Summary of a linear mixed-effects model comparing effects of ambient versus structural controls on daily soil moisture (% volumetric water content, VWC) in climate change experiments across the year.** Estimates (est.) are the intercept (representing ambient controls) and coefficient (representing structure effects) from the models; se is the standard error for these estimates. For this annual model, the difference between control types was significant based on Type II Wald χ^2 tests of fixed effects (χ^2=89.95, df=1, p<0.001). Random effects of site (n=5) and year nested within site (n=21 year-site combinations) were fit with a random slope and intercept structure; total number of observations= 44,468.

	soil moisture (vwc)	
predictor	est.	se
intercept	21.20	1.86
structure effect	-2.43	0.26

Table S9: **Summaries of linear mixed-effects models, fit to each month of data, comparing effects of ambient versus structural controls on daily mean, minimum, and maximum soil temperature, fit to each month separately**, consistent with Figure 4 in the main text. Estimates (est.) are the intercept (representing ambient controls) and coefficient (representing structural effects) from the models; se is the standard error for these estimates. We list test statistics, and p-values for Type II Wald χ^2 tests of fixed effects (df=1 for all tests). Random effects of site (n=5 for all mean soil temperature models; n=4 for all min and max soil temperature models) and year nested within site (n=19 or 20 year-site combinations for all mean soil temperature models; n=18 or 19 for all min and max soil temperature models) were fit with a random slope and intercept structure; total number of observations ranged from 3,814 to 4,186; units are °C.

		mean soil temp.				min soil temp.				max soil temp.			
mon	predictor	est.	se	χ^2	p	est.	se	χ^2	p	est.	se	χ^2	p
01	intercept	2.66	1.25	3.63	0.057	2.34	1.21	2.09	0.149	3.92	1.65	13.71	<0.001
	structure effect	-0.45	0.23			-0.72	0.50			-0.35	0.09		
02	intercept	2.86	1.44	13.06	<0.001	2.58	1.26	3.24	0.072	4.66	1.92	1.99	0.158
	structure effect	-0.44	0.12			-0.67	0.37			-0.41	0.29		
03	intercept	5.24	1.78	6.44	0.011	4.66	1.58	3.64	0.056	7.75	2.04	0.92	0.337
	structure effect	-0.44	0.17			-0.44	0.23			-0.50	0.52		
04	intercept	9.98	1.85	8.53	0.003	8.93	1.98	10.52	0.001	13.24	1.80	0.96	0.327
	structure effect	-0.67	0.23			-0.65	0.20			-0.63	0.65		
05	intercept	14.92	1.37	3.85	0.05	13.74	1.54	4.91	0.027	17.54	1.41	0.59	0.441
	structure effect	-0.31	0.16			-0.27	0.12			-0.32	0.42		
06	intercept	18.29	1.58	0	0.972	17.43	1.57	0.76	0.383	20.98	1.78	0.04	0.844
	structure effect	-0.01	0.20			-0.14	0.16			0.09	0.47		
07	intercept	21.07	1.33	0.06	0.815	19.97	1.34	0.45	0.501	23.76	1.46	0.01	0.914
	structure effect	-0.07	0.28			-0.12	0.18			-0.07	0.61		
08	intercept	20.93	1.20	2.56	0.11	19.59	1.29	1.35	0.244	23.23	1.42	1.58	0.209
	structure effect	-0.26	0.16			-0.20	0.17			-0.37	0.30		
09	intercept	18.23	1.24	10.15	0.001	16.94	1.36	0.58	0.445	20.54	1.43	1.74	0.188
	structure effect	-0.36	0.11			-0.21	0.27			-0.40	0.31		
10	intercept	13.03	1.22	10.48	0.001	12.26	1.24	1.39	0.239	15.42	1.39	10.02	0.002
	structure effect	-0.42	0.13			-0.56	0.48			-0.50	0.16		
11	intercept	8.27	1.13	1.87	0.172	7.34	1.23	0.83	0.363	10.11	1.43	3.16	0.075
	structure effect	-0.33	0.24			-0.52	0.57			-0.28	0.16		
12	intercept	5.03	1.21	2.8	0.094	4.38	1.24	1.53	0.215	6.40	1.53	4.83	0.028
	structure effect	-0.40	0.24			-0.61	0.49			-0.26	0.12		

Table S10: **Summaries of linear mixed-effects models, fit to each month comparing effects of ambient versus structural controls on daily minimum and maximum above-ground temperature, fit to each month separately**, consistent with Figure 4 in the main text. Estimates (est.) are the intercept (representing ambient controls) and coefficient (representing structural effects) from the models; se is the standard error for these estimates. We list test statistics, and p-values for Type II Wald χ^2 tests of fixed effects (df=1 for all tests). Random effects of site (n=4 for both models) and year nested within site (n=18 year-site combinations for both models) were fit with a random slope and intercept structure; total number of observations was 3,726; units are °C.

		min air temp.				max air temp.			
mon	predictor	est.	se	χ^2	p	est.	se	χ^2	p
01	intercept	-5.49	1.78	5.27	0.022	5.09	2.60	0.01	0.927
	structure effect	0.61	0.26			-0.03	0.29		
02	intercept	-3.92	1.83	1.41	0.235	7.10	3.03	2.93	0.087
	structure effect	0.55	0.46			0.36	0.21		
03	intercept	-0.08	1.55	8.59	0.003	12.60	2.41	2.75	0.097
	structure effect	0.50	0.17			0.52	0.31		
04	intercept	5.28	1.80	9.33	0.002	19.27	1.92	18.31	<0.001
	structure effect	0.55	0.18			1.26	0.29		
05	intercept	11.62	1.46	6.56	0.01	23.49	1.03	7.75	0.005
	structure effect	0.48	0.19			0.77	0.28		
06	intercept	15.45	1.47	10.13	0.001	26.32	1.82	4.4	0.036
	structure effect	0.43	0.14			0.59	0.28		
07	intercept	17.90	1.26	4.47	0.035	28.94	1.25	3.58	0.059
	structure effect	0.85	0.40			0.61	0.32		
08	intercept	17.07	1.43	2.07	0.15	27.39	1.15	0.87	0.35
	structure effect	0.65	0.45			0.33	0.35		
09	intercept	13.34	1.39	4.71	0.03	23.72	1.47	2.66	0.103
	structure effect	0.88	0.41			0.38	0.23		
10	intercept	7.26	1.26	4.27	0.039	17.29	1.70	1.89	0.169
	structure effect	0.79	0.38			0.30	0.22		
11	intercept	1.21	1.25	4.23	0.04	12.79	1.83	2.76	0.097
	structure effect	0.88	0.43			0.26	0.15		
12	intercept	-2.83	1.48	5.29	0.021	7.56	2.38	0.26	0.61
	structure effect	0.43	0.19			-0.11	0.23		

Table S11: **Summaries of linear mixed-effects models, fit to each month comparing effects of ambient versus structural controls on soil moisture (% VWC), fit to each month separately**, consistent with Figure 4 in the main text. Estimates (est.) are the intercept (representing ambient controls) and coefficient (representing structural effects) from the models; se is the standard error for these estimates. We list test statistics, and p-values for Type II Wald χ^2 tests of fixed effects; df=1 for all models. Random effects of site (n=4) and year nested within site (n=18 year-site combinations) were fit with a random slope and intercept structure; total number of observations was 3,829.

mon	predictor	est.	se	χ^2	p
01	intercept	22.58	3.23	59.24	<0.001
	structure effect	-2.77	0.36		
02	intercept	22.10	3.24	16.78	<0.001
	structure effect	-2.54	0.62		
03	intercept	23.58	2.43	8.3	0.004
	structure effect	-2.48	0.86		
04	intercept	22.54	2.15	9.24	0.0024
	structure effect	-2.06	0.68		
05	intercept	21.08	2.31	40.17	<0.001
	structure effect	-2.20	0.35		
06	intercept	18.44	1.37	30.78	<0.001
	structure effect	-2.12	0.38		
07	intercept	17.60	2.18	20.22	<0.001
	structure effect	-2.38	0.53		
08	intercept	16.59	1.90	12.95	<0.001
	structure effect	-2.09	0.58		
09	intercept	15.99	1.54	13.2	<0.001
	structure effect	-1.79	0.49		
10	intercept	20.15	1.93	20.9	<0.001
	structure effect	-2.27	0.50		
11	intercept	21.18	1.77	21.9	<0.001
	structure effect	-2.70	0.58		
12	intercept	22.74	2.83	15.64	<0.001
	structure effect	-2.88	0.73		

Table S12: **Comparison of estimated warming (Tdiff) attained in infrared studies, per degree of target warming, with different types of controls**, structural versus ambient . Tdiff was calculated by substracting mean annual temperature in control plots from mean annual temperature each treatment level among controls, and in seven infrared studies that measured above-ground warming (including canopy, surface, and air temperature). We divided Tdiff by the target warming for the treatment level, to standardize across studies. We then fit a mixed-effects model with Tdiff as the response, control type (structural or ambient) as the predictor, and random effects of site (n=7) and year nested within site (24 year:site combinations), with an intercept-only structure. We list the estimated Tdiff and its standard error (se) for each control type. Tdiff differed significantly across the two control types, based on Type II Wald χ^2 tests of the fixed effects (χ^2=19.71 , df=1, p<0.001). Total number of observations was 37.

	Tdiff	se
structural controls	1.01	0.09
ambient controls	-0.59	0.13

Table S13: **Summary of a linear mixed-effects models of how precipitation treatment affects temperature in climate change experiments.** We include data from all studies that manipulated precipitation and measured daily above-ground temperature and/or soil temperature. Estimates (est.) are the intercept and coefficients for precipitation and warming treatments, as well as their interaction, from the model; se is the standard error for these estimates; p-values represent significance tests for Type III Wald χ^2 tests. Random effects were site (n=4 for above-ground, n=3 for soil), year of study (n=10 year:site combinations for above-ground, n=9 for soil), and doy nested within year (n=2818 doy:year:site combinations for above-ground, n= 2747 for soil), with a random intercept structure. Total number of observations was 78347 for above-ground temperature and 76481 for soil.

response	predictors	est.	se	χ^2	df	p
min above-ground temp.	intercept	6.68	0.75	78.59	1	<0.001
	preciptreat	-0.01	0.00	930.79	1	<0.001
	target	0.74	0.01	4391.28	1	<0.001
	precip*target	0.00	0.00	62.73	1	<0.001
max above-ground temp.	intercept	23.80	1.09	477.23	1	<0.001
	preciptreat	-0.02	0.00	2384.33	1	<0.001
	target	0.78	0.02	1819.10	1	<0.001
	precip*target	-0.00	0.00	31.73	1	<0.001
min soil temp.	intercept	18.27	0.97	356.27	1	<0.001
	preciptreat	-0.01	0.00	1381.13	1	<0.001
	target	1.26	0.02	5696.28	1	<0.001
	precip*target	-0.00	0.00	720.76	1	<0.001
max soil temp.	intercept	12.15	0.62	378.95	1	<0.001
	preciptreat	-0.01	0.00	2280.12	1	<0.001
	target	0.71	0.01	5112.44	1	<0.001
	precip*target	0.00	0.00	50.49	1	<0.001

Table S14: **Summary of a linear mixed-effects model of how target warming treatment affects soil moisture (% VWC) in climate change experiments with different types of active warming.** We excluded data from plots with precipitation treatments from this analysis. Estimates (est.) are the coefficients for warming type and the interaction between warming type and target warming (i.e. the type-specific effect of target warming) from the model; se is the standard error for these estimates. The effects of warming type was significant (χ^2=170.44, df=3, p<0.001), as was target warming (χ^2=4271.77, df=1, p<0.001), and the interaction of warming type and target warming (χ^2=552.89, df=2, p<0.001), based on Type II Wald χ^2 tests of the fixed effects. Random effects were site (n=12), year of study nested within site (45 year-site combinations), and doy nested within year (n=8897 doy-year-site combinations), with a random intercept structure. Total number of observations was 136,041.

	est.	se
infrared	21.00	1.78
forced air	21.51	4.87
forced air and soil	17.22	2.80
target (infrared)	-0.80	0.02
target (forced air)	-0.50	0.03
target(forced air and soil)	-0.30	0.02

Table S15: **Summary of a linear mixed-effects model comparing soil moisture (% VWC)** in climate change experiments with different types of active warming and in experimentally warmed plots to two different control types, structural and ambient controls. We excluded data from plots with precipitation treatments from this analysis. Estimates (est.) are the intercept (representing mean moisture in ambient controls) and coefficients from the from the model (i.e. differences between the ambient) for structural controls and warmed plots (pooled across all target warming levels); se is the standard error for these estimates. There were significant differences among warming types based on Type II Wald χ^2 tests of the fixed effect (χ^2=7229.01, df=2, p<0.001). Random effects were site (n=10), year nested within site (n=35 site-year combinations), and doy nested within year (7,979 doy-year-site combinations) with a random intercept structure. Total number of observations was 72730.

	est.	se
infrared	19.79	1.89
forced air and soil	-1.03	3.53
infrared:structure effect	1.43	0.16
infrared:warmed effect	-0.50	0.15
forced air and soil:structure effect	-3.89	0.16
forced air and soil:warmed effect	-2.24	0.15

Table S16: **Comparison of linear mixed-effects models for budburst day of year** that contain either target warming treatment only as a fixed effect or measured mean annual above-ground temperature (°C), mean soil moisture from January through March (in % VWC), and their interaction as fixed effects. Estimates (est.) are the intercept and coefficients from the models; se is the standard error for these estimates. Both models include random effects of site (n=5), year of study nested within site (n=13 year-site combinations), and plant species (n=54), each with a random intercept structure. Total number of observations was 12,549. Analysis includes all stuidies that monitored budburst, and measured soil moisture and above-ground temperature (exp01,exp03,exp04,exp07,exp10)

V1	predictor	est.	se
target warming model	intercept	110.46	6.72
	target	-2.09	0.08
tmin*soilmois model	intercept	150.69	4.48
	tmin	-6.04	0.36
	soilmois_janmar	-1.35	0.12
	tmin*soilmois_janmar	0.15	0.02

Table S17: **Comparison of budburst model fits** from four models with different fixed effects: 1) target warming only, 2) measured mean annual above-ground temperature (tmean) only, tmean and soil moisture (soilmois), and tmean, soilmois, and their interaction. We additionally compared models with mean annual maximum temperature, and seasonal temperature variables; we present only tmean here because this variable provided the best model fit (i.e. lowest AIC). Models that included both measured temperature and soil moisture provided better fit than the target warming model. The target warming model provided better fit than measured temperature alone. This is consistent with our findings that experimental warming changes more than temperature alone, and further suggests that the changes beyond temperature are biologically important.

model	df	AIC	ΔAIC
tmean*soilmois	8	107225.81	0.00
tmean+soilmois	7	107309.38	83.58
target	6	107346.19	120.39
tmean	6	107396.44	170.64

Supplemental Figures

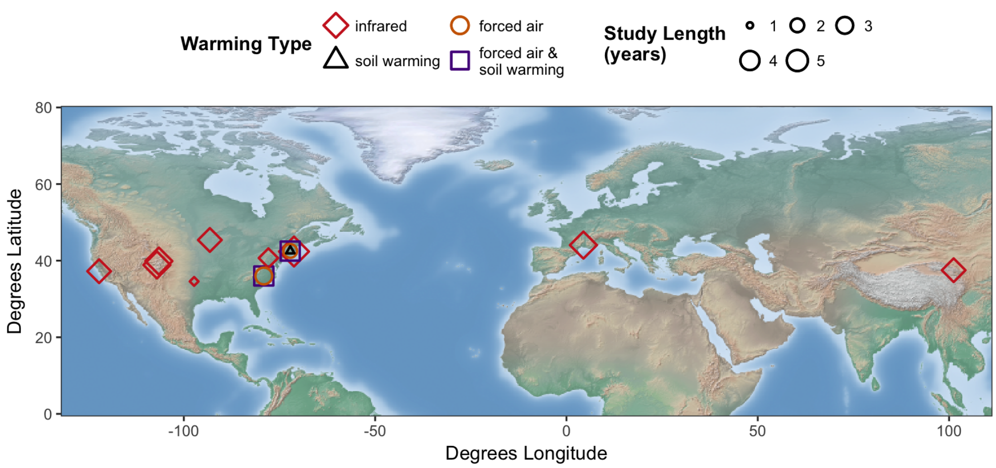

Figure S1: **Climate data from 12 climate change experiments in North America and Europe are included in the C3E database and analyzed here.** See Tables S1 and S2 for details.

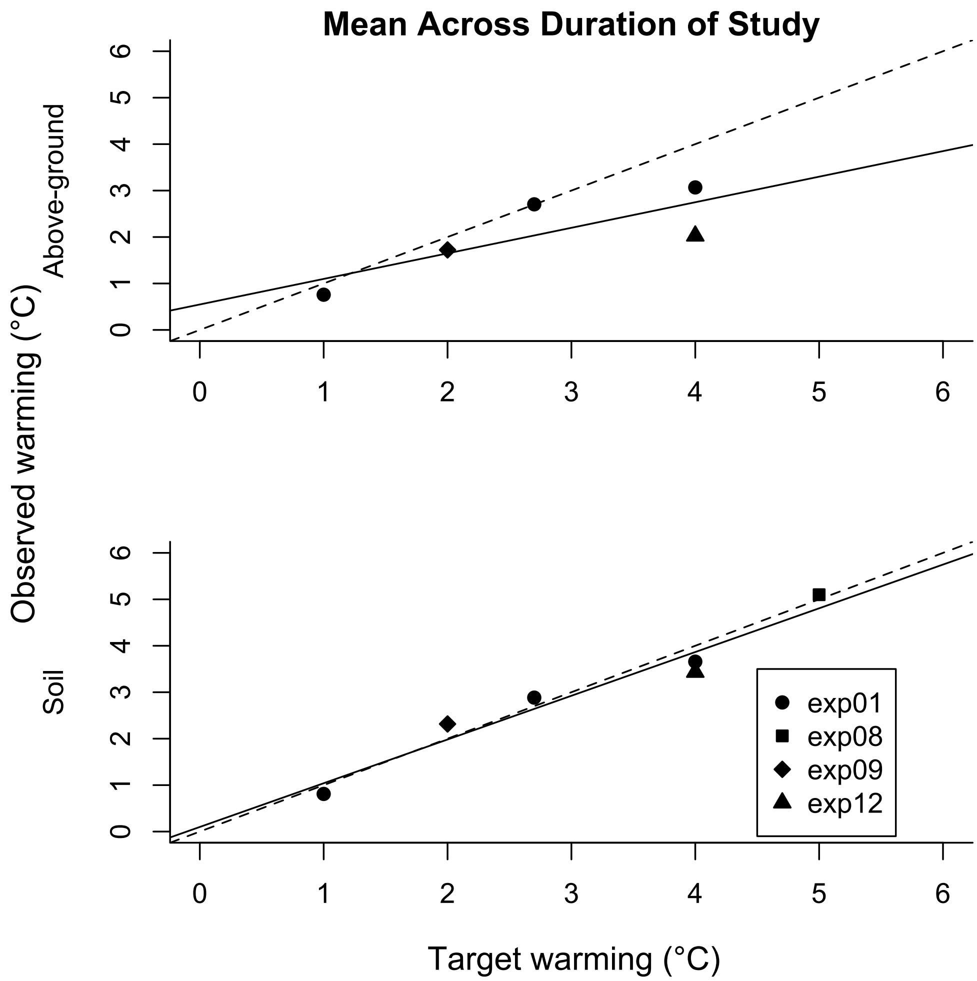

Figure S2: **Mean warming across the study duration.** Reporting only the mean temperature difference across the duration of the study, as is commonly done in publications of climate change experiments, hides potentially important variations in daily, seasonal, and annual temperatures among treatments. The solid line is the fitted relationship between observed and target warming and the dashed line shows when observed warming is exactly equal to target warming (1:1). Compare to Figure 2 in the main text. (Note that exp01 did not specify explicit target temperatures for all warming treatments. See **??** for comparison using reported, rather than target, temperatures.)