How do climate change experiments alter local climate?

A. Ettinger[1,2,a], I. Chuine[3,b], B.I. Cook[4,5,c], J.S. Dukes[6,d], A.M. Ellison[7,e], M.R. Johnston[8,f], A.M. Panetta[9,g], C.R. Rollinson[10,h], Y. Vitasse[11,12,i], and E.M. Wolkovich[1,8,j]

[1]Arnold Arboretum of Harvard University, Boston, Massachusetts 02131, USA

[2]Tufts University, Medford, Massachusetts 02155, USA

[3]CEFE UMR 5175, CNRS, Université de Montpellier,Université Paul-Valéry Montpellier, EPHE IRD, Montpellier, France

[4]Lamont-Doherty Earth Observatory, Columbia University, Palisades, New York 10964, USA

[5]NASA Goddard Institute for Space Studies, New York, New York 10025, USA

[6]Department of Forestry and Natural Resources and Department of Biological Sciences, Purdue University, West Lafayette, Indiana 47907, USA

[7]Harvard Forest, Harvard University, Petersham, Massachusetts 01366, USA

[8]Department of Organismic and Evolutionary Biology, Harvard University, Cambridge, Massachusetts 02138, USA

[9]Department of Ecology and Evolutionary Biology, University of Colorado, Boulder, Colorado 80309, USA

[10]The Morton Arboretum, Lisle, Illinois 60532, USA

[11]Institute of Geography, University of Neuchâtel, Neuchâtel, Switzerland

[12]Swiss Federal Institute for Forest, Snow and Landscape Research WSL, Neuchâtel, Switzerland

[a]Corresponding author; email: aettinger@fas.harvard.edu; phone: 781-296-4821; mailing address: 1300 Centre Street, Boston, Massachusetts 02140, USA

[b]isabelle.chuine@cefe.cnrs.fr

[c]bc9z@ldeo.columbia.edu

[d]jsdukes@purdue.edu

[e]aellison@fas.harvard.edu

[f]mjohnston@g.harvard.edu

[g]anne.panetta@colorado.edu

[h]crollinson@mortonarb.org

[i]yann.vitasse@wsl.ch

[j]wolkovich@fas.harvard.edu

April 11, 2018

Statement of authorship All authors conceived of this manuscript, which was inspired by our discussions at a Radcliffe Exploratory Seminar in 2016, and all authors contributed to manuscript revisions. A.E. and E.W. conceived of the idea for the literature review, database compilation, and related Radcliffe Exploratory Seminar. A.E. compiled the datasets; A.E. and C.R. analyzed the data and created the figures; A.E. wrote the manuscript.

Data Accessibility The MC3E database will be available at KNB (Ettinger & Wolkovich, 2018), along with all R code from the analyses included in this paper. (Currently, metadata are published there; the full database and R code are available to reviewers on github.)

Running title Experimental climate change

Key words global warming, warming experiment, microclimate, soil moisture, spring phenology, budburst, direct and indirect effects, structural control, hidden treatment, active warming, target temperature, feedback

Type of article Review and Synthesis

Number of words in abstract 200

Number of words in main text 5,465

Number of references 87

Number of figures 5

Number of tables 0

Number of text boxes 1

Number of words in Box 1 475

Abstract

To understand and forecast biological responses to climate change, scientists frequently use field experiments that alter temperature and precipitation. Climate manipulations can manifest in complex ways, however, challenging interpretations of biological responses. We reviewed publications on active warming experiments to compile a database of daily plot-level climate data from 15 experiments that use forced air, infrared heaters, and soil cables to warm plots in forests, alpine meadows, and grasslands. We find that the common practices of analyzing primarily mean changes among treatments and analyzing treatments as categorical variables (e.g. warmed verses unwarmed) masks important variation in treatment effects over space and time. Measured mean warming in plots with the same target warming can vary by 3°C or more among blocks. Furthermore, warming treatments produce secondary effects, such as soil drying. The implications of these complexities can have important biological consequences. We show one such consequence with a case study of plant phenology, in which accounting for secondary non-temperature effects of warming triples the estimated sensitivity of budburst to warming. Based on our synthesis, we present recommendations for future analyses, experimental design and data sharing that will improve the ability of climate change experiments to accurately identify and forecast species' responses.

Introduction

Climate change is dramatically altering earth's biota, shifting the physiology, distribution, and abundance of organisms, with cascading community, ecosystem, and climate effects (Shukla & Mintz, 1982; Cox *et al.*, 2000; Thomas *et al.*, 2004; Parmesan, 2006; Field *et al.*, 2007; Sheldon *et al.*, 2011; Urban *et al.*, 2012). Much uncertainty exists about how particular individuals, populations, species, communities, and ecosystems will respond as shifts in temperature and precipitation regimes become more extreme (Thuiller, 2004; Friedlingstein *et al.*, 2014). Predicting biological responses to current and future climate change—and their feedbacks to earth's climate and ecosystem services—are among the most significant challenges facing ecologists today.

Two common approaches for understanding biological effects of climate change are observational studies and process-based modeling; yet these approaches are insufficient for several reasons. Observational studies, which correlate recorded biological patterns with measured trends in climate, cannot disentangle the causal effects of warming from other factors that have also changed over time, such as successional stage or land use. Models based on correlative data may fail to make useful predictions for future conditions that fall outside the range of historical variability (e.g., Pearson & Dawson, 2004; Hampe, 2004; Ibanez *et al.*, 2006; Swab *et al.*, 2012; Chuine *et al.*, 2016). Climate change will yield warmer temperatures than the previous 150 years, and possibly warmer than at any time in the last 2000 years (Ohlemüller *et al.*, 2006; Williams & Jackson, 2007; Williams *et al.*, 2007; Stocker *et al.*, 2013). Process-based models begin to overcome these challenges through inclusion of explicit mechanistic relationships between climate and biological outcomes. However, they are limited by the processes they include (i.e., our understanding of mechanism), as well as by the data available to parameterize those processes (Moorcroft, 2006; Kearney & Porter, 2009).

Experimental data from field-based climate change experiments are crucially important to fill these knowledge gaps and determine mechanistic links between climate change and biological responses. Experiments can quantify biological responses to different levels of climate change, and can create the "no-analog" climate scenarios forecasted for the future, particularly when they employ active warming methods, such as forced air heaters, soil warming cables, or infrared heaters (Shaver *et al.*, 2000; Williams *et al.*, 2007; Aronson & McNulty, 2009). In addition, active warming can be combined with precipitation manipulations (e.g., snow removal, water additions or reductions), offering the ability to isolate effects of temperature and precipitation from other environmental changes (e.g., Price & Waser, 1998; Cleland *et al.*, 2006; Sherry *et al.*, 2007; Rollinson & Kaye, 2012). Compared with indoor growth-chamber experiments, field-based experiments offer the possibility of preserving important but unknown or unquantified feedbacks among biotic and abiotic components of the studied systems.

With climate change experiments, ecologists often aim to test hypotheses about how projected warming will affect species' growth, survival, and future distributions (Dukes & Mooney, 1999; Hobbie *et al.*, 1999; Morin *et al.*, 2010; Pelini *et al.*, 2011; Chuine *et al.*, 2012; Reich *et al.*, 2015; Gruner *et al.*, 2017). But is it reasonable to extrapolate findings from these experiments to the real world? In what ways is microclimate altered by climate manipulations? Recent research suggests that climate manipulations may not always alter microclimate in ways that are consistent with observed changes over time (Wolkovich *et al.*, 2012; Menke *et al.*, 2014; Andresen *et al.*, 2016). We need detailed assessments of how active warming experiments alter the climate conditions experienced by organisms, and the extent to which these conditions are similar to current field conditions or anticipated climate change.

Here, we investigate the complex ways that plot-level climate is altered by active-warming treatments, both directly and indirectly, across multiple studies. The qualitative challenges and opportunities of climate change experiments have been summarized previously (e.g., De Boeck *et al.*, 2015) and effects of these manipulations on some aspects of microclimate have been published for individual sites (e.g., Harte *et al.*, 1995; McDaniel *et al.*, 2014b; Pelini *et al.*, 2011). However, an in-depth quantitative meta-analysis is lacking. Using plot-level daily microclimate data from 15 active warming experiments (yielding 59 experiment years and 14,913 experiment days; Table S1), we show the direct and indirect ways that experimental manipulations alter microclimate. We use a case study of spring plant phenology to demonstrate how common analyses that assume a constant warming effect and ignore secondary effects of warming treatments as drivers of biological responses lead to inaccurate quantification of plant sensitivity to temperature shifts. Finally, we synthesize our findings to make recommendations for future analysis and design of climate change experiments (Box 1).

MicroClimate from Climate Change Experiments (MC3E) database

To investigate how climate change experiments alter microclimate, we first identified published, active-warming field experiments. We focused on *in situ* active-warming manipulations because recent analyses indicate that active-warming methods are the most controlled and consistent methods available for experimental warming (Kimball, 2005; Kimball *et al.*, 2008; Aronson & McNulty, 2009; Wolkovich *et al.*, 2012). We do not include passive warming experiments because they have been analyzed extensively already and are known to have distinct issues, including extreme reduction in wind, overheating and great variation in the amount of warming depending on irradiance and snow depth (Marion *et al.*, 1997; Shaver *et al.*, 2000; Wolkovich *et al.*, 2012; Bokhorst *et al.*, 2013). (We do include a summary comparison of warming data from passive open top chambers, taken from Bokhorst *et al.* (2013) with warming data from different active warming methods in Table S2.)

We carried out a full literature review to identify potential active warming field experiments to include in the database. To find these studies, we followed the methods and search terms of Wolkovich *et al.* (2012) for their Synthesis of Timings Observed in iNcrease Experiments (STONE) database (Wolkovich *et al.*, 2012), but restricted our focus to active-warming experiments. Further, because our goal was to tease out variation in climate (including temperature and soil moisture), we focused on warming studies with multiple levels of warming and/or precipitation treatments. These additional restrictions constrained the list to 11 new studies published after the STONE database, as well as six of the 37 studies in the STONE database. We contacted authors to obtain daily climate data and phenological data for these 17 studies and received data (or it was already publicly available) for 10 of them, as well as datasets from five additional sites offered or suggested to us over the course of our literature review and data analysis. The daily temperature and soil moisture data from these 15 experiments were put together into the MicroClimate from Climate Change Experiments (MC3E) database (Figure S1, Table S1), which is available at KNB (Ettinger & Wolkovich, 2018).

Complexities in interpreting experimental climate change

Climate change experiments often include detailed monitoring of climate variables at the plot level, yielding large amounts of data, such as daily or hourly temperature and other climate variables, over the course of an experiment. Ecologists, however, are generally interested in the ecological responses (e.g., community dynamics, species' growth, abundance, or phenology), which are collected on much coarser timescales (e.g., weekly or annually). Not surprisingly, then, when analyzing ecological responses, authors typically provide detailed information on the observed biological responses, and report only the mean change in climate over the course of the experiment and whether it matched their target level of change (e.g., Price & Waser, 1998; Rollinson & Kaye, 2012; Clark *et al.*, 2014a,b). Several studies have conducted detailed, independent analyses of microclimate data from warming experiments (e.g., Harte *et al.*, 1995; Kimball, 2005; Kimball *et al.*, 2008; McDaniel *et al.*, 2014b; Pelini *et al.*, 2011). These detailed analyses provide valuable case studies of experimental effects on microclimate data alone, but have generally not been incorporated into analyses of ecological responses.

Though the focus in interpreting ecological responses to climate change manipulations has been primarily on mean shifts in microclimate, the imposed manipulations result in much more complex shifts. The magnitude of change in these manipulations varies in time and space, and the presence of experimental equipment alone (with no heat added) often alters environmental conditions. These factors, discussed below, challenge our interpretation of how experimental warming studies forecast effects of climate change on organisms and ecosystems.

Effects on local climate vary over time and space

Reporting only the mean temperature difference across the duration of the study hides potentially important variations in daily, seasonal, or annual temperatures among treatments (compare Figure 2 to Figure S2). Using the MC3E database, we found that active warming reduces above-ground daily temperature range by 0.37°C per °C of target warming (Table S3, see also Table S1, which details the different methods used to measure temperature). Active warming decreased above-ground daily temperature range by differentially affecting maximum and minimum temperatures: warming increased daily minima by 0.81°C per °C of target warming, but only increased daily maxima by 0.48°C per °C of target warming (Table S3). These effects varied by site (Table S3), but we found no clear patterns by warming type (e.g., infrared versus forced air). Soil daily temperature range was not affected by experimental warming (not shown).

We observed strong seasonal and annual variations in experimental warming effects (Figures 1, 2, Table S4). Warming appears to be generally close to targets in winter and early spring, and farthest below targets in summer (day of year150-200), though patterns differ among sites (Figures 1). The variation in warming effectiveness may be driven by interactions between warming treatments and daily, seasonal, and annual weather patterns, since the magnitude of warming can vary as weather conditions change. Both infrared heaters and soil cables fail to achieve the target temperature increases during rainstorms (Peterjohn *et al.*, 1993; Hoeppner & Dukes, 2012) and with windy conditions (Kimball, 2005; Kimball *et al.*, 2008). In addition, treatments are often applied inconsistently within or across years. Heat applications are frequently shut off during winter months, and some heating methods, even if left on throughout the year, are not capable of applying constant warming year-round (e.g. Clark *et al.*, 2014a,b; Hagedorn *et al.*, 2010).

Treatment effects also vary spatially, adding further complication to interpreting effects of climate change experiments. The MC3E database contains six studies that used blocked designs, allowing us to examine spatial variation in the amount of warming (i.e. the difference between treatment and control plots within a block). These studies include five infrared and one soil warming experiment, all of which used structural controls only. We found that the amount of observed warming frequently varied by more than 1°C (and up to 3°C) among blocks (Figure 2, Table S5); this block-to-block variation in warming treatment is significant, at 60-100% of target temperatures. These differences in warming among blocks may be caused by fine-scale variation in vegetation, slope, aspect, soil type, or other factors that can alter wind or soil moisture, which in turn affect warming (Peterjohn *et al.*, 1993; Kimball, 2005; Kimball *et al.*, 2008; Hoeppner & Dukes, 2012;

Rollinson & Kaye, 2015).

Of course, identical experimental treatments across space and time are not necessary, nor realistic, for robust analysis of experimental results or for forecasting. Indeed, the spatial and temporal variation we report could improve and refine models, and—at least in some regions—may be consistent with contemporary patterns of climate change (Stocker *et al.*, 2013). Taking advantage of this variation, however, requires understanding and reporting it (e.g., Milcu *et al.*, 2016). In contrast, fine-scale spatial and temporal variations in warming treatments are rarely analyzed explicitly with ecological data, so the implications for interpretation of experimental findings are unclear.

Experimental infrastructure alters local climate

Experimental structures themselves can alter temperature and other important biotic and abiotic variables in ways that are not generally examined in experimental climate change studies. The importance of controls that mimic a treatment procedure without actually applying the treatment is widely acknowledged in biology (e.g., Spector, 2001; Johnson & Besselsen, 2002; Quinn & Keough, 2002). Though some studies include treatments with non-functional warming equipment as well as ambient controls in experimental climate change studies, the magnitude and implications of structural effects on climate are rarely interpreted or analyzed.

To investigate the magnitude of infrastructure effects, we compared temperature and soil moisture data from five active warming studies at two sites: Duke Forest and Harvard Forest (Farnsworth *et al.*, 1995; Clark *et al.*, 2014b; Marchin *et al.*, 2015; Pelini *et al.*, 2011). These were the only studies in the MC3E 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 (n=1), forced air chambers (n=2), or both (n=2), but with no heat applied) and ambient controls with no infrastructure added. Other studies, all of which utilized infrared warming, monitored environmental conditions in only structural controls (n=5) or only ambient controls (n=4). (One study, exp15, which utilized forced air, monitored environmental conditions in both ambient and structural controls, but we were only able to obtain data for the ambient controls so it is excluded from this analysis.)

We found that experimental structures altered above-ground and soil temperatures in opposing ways: above-ground temperatures were higher in the structural controls than in ambient controls, whereas soil temperatures were lower in structural controls compared with ambient controls (Figure 3a-d). This general pattern was consistent across different temperature models (mean, minimum, and maximum temperatures), although the magnitude varied among seasons, studies, and years (Figure 3a-d, Tables S6, S7, S9, S10). We also found that experimental infrastructure decreased soil moisture relative to ambient conditions across all seasons, studies, and years (Figure 3e, Tables S8, S11).

There are several possible reasons for the observed climatic differences between ambient and structural controls. Infrastructure materials may shade the plots, reduce airflow, reduce albedo relative to surroundings, or otherwise change the energy balance. Specifically, soil temperatures may be cooler in structural controls because the experimental structures block sunlight from hitting the ground surface, which would therefore experience less radiative heating than ambient controls. In addition, above-ground temperatures may be warmer in structural controls because the structures radiatively warm the air around them and block wind, inhibiting mixing with air outside of the plot. Structures also interfere with precipitation hitting the ground, thereby reducing local soil moisture and snowpack, with its insulative properties. The latter likely plays a bigger role in soil temperature differences at the Harvard Forest sites (exp04, exp07, exp08), where average annual snowfall is over one meter, than at Duke Forest (exp03, exp10), where average snow accumulation each winter is 20 cm or less. Finally, for some warming types (e.g., soil cables), structural controls experience increased soil disturbance compared with ambient controls; this may alter water flow and percolation, and introduce conductive material such as metal via the cables or posts.

We were unable to compare ambient and structural controls for experiments using infrared heating, because no studies in our database included both control types. Of the seven infrared sites that measured soil temperature, three used ambient controls (exp06, exp11,exp14) and four used structural controls (exp01, exp09, exp12, exp13). Thus, we were able to compare differences between the two control types and the measured amount of warming per degree of target warming. We expected that, if infraread heating infrastructure affects microclimate when no heat is applied, the measured amount of warming per degree of target warming should differ by control type. Among these studies, warming per °C of target warming was close to one for structural controls (1.01°C, SE= 0.09, Table S12), but was significantly lower for ambient controls (0.41 °C of warming per °C of target warming, SE=0.09, Table S12). The number of studies using ambient controls is small (n=3) and includes only two different sites, so we hesitate to draw strong conclusions. The trend is suggestive, though, that infrared heating equipment alone alters microclimate, likely by shading the soil surface (McDaniel *et al.*, 2014b).

Although there is little discussion of measured temperature (or other) differences between ambient and structural control plots in published work (e.g., Farnsworth *et al.*, 1995; Pelini *et al.*, 2011; Clark *et al.*, 2014a), the few studies that do mention these differences are consistent with our findings. Clark *et al.* (2014b), whose study employed forced air and soil cables for warming, state that "control of the air temperature was less precise, in part due to air scooping on windy days." Marchin *et al.* (2015), who used forced air warming, note that structural controls had mean spring air temperatures about 0.5°C or more above ambient temperatures. Peterjohn *et al.* (1994), who warmed soil with heating cables, reported cooler soil temperatures in structural controls than in ambient controls at shallow soil depths. Similarly, we found the greatest difference in soil temperature between structural and ambient controls in shallow soils (e.g., exp10, soil depth = 2cm). If addressed, the focus to date has been largely on these abiotic impacts of experimental structures, but structures also alter herbivory and other biotic conditions (Kennedy, 1995; Moise & Henry, 2010; Wolkovich *et al.*, 2012; Hoeppner & Dukes, 2012).

Most warming experiments calculate focal response variables relative to ambient controls (e.g., Price & Waser, 1998; Dunne *et al.*, 2003; Cleland *et al.*, 2006; Morin *et al.*, 2010; Marchin *et al.*, 2015), which our analyses suggest will not properly account for infrastructure effects. In addition, results from studies reporting only structural controls (e.g., Sherry *et al.*, 2007; Hoeppner & Dukes, 2012; Rollinson & Kaye, 2012), should be cautiously applied outside of an experimental context, as—without ambient controls—their inference is limited to the environment of the structural controls. Though a major additional effort, our results suggest that studies aiming to predict or forecast effects at their particular location would benefit markedly from employing both structural and ambient controls. This will allow for documentation and analysis of infrastructure effects on abiotic and biotic responses. Separating infrastructure artifacts from warming effects is critical if we wish to apply findings to forecasts outside of an experimental context.

Secondary and feedback effects of climate change manipulations

Climate change experiments often seek to manipulate temperature or precipitation separately as well as interactively, but manipulating either of these variables in isolation is difficult. Treatments involving precipitation additions typically reduce temperatures in climate change manipulations (Sherry *et al.*, 2007; Rollinson & Kaye, 2012; McDaniel *et al.*, 2014b). For example, McDaniel et al. (2014) observed that a 20% increase in precipitation reduced mean hourly temperatures by 0.3°C over the course of their two-year experiment.

In the MC3E database, there are four experiments that manipulated both temperature and precipitation, and provided daily above-ground temperature data (three of these also measured soil temperature). Across these studies, all of which used infrared heating, we found that increasing the amount of added precipitation reduced daily minimum and maximum above-ground temperatures, at rates of 0.01 and 0.02 °C, respectively, and soil temperatures, at a rate of 0.01°C for both minimum and maximum temperature, per percent increase in added precipitation (Table S13). This is because increasing soil moisture (an effect of precipitation additions) typically shifts the surface energy balance to favor latent (i.e., evapotranspiration) over sensible energy fluxes, reducing heating of the air overlying the soils. Three of the four studies in this analysis (exp01, exp05, exp12) used a constant wattage output from their infrared heaters; one (exp09) used an independent feedback approach, in which temperature measurements are used to vary the voltage input. It may be that feedback approaches are able to maintain higher warming levels across precipitation additions. However, maintaining target warming levels is a challenge for feedback systems as well, particularly during seasons or years with wetter soils and higher evapotranspiration (Rich *et al.*, 2015), and it is unclear how either approach compares to conditions outside of an experimental context.

Experimental warming generally increases vapor pressure deficit and reduces soil water content (e.g., Harte *et al.*, 1995; Sherry *et al.*, 2007; Morin *et al.*, 2010; Pelini *et al.*, 2014; Templer *et al.*, 2016). Of the 15 experiments in the MC3E database, we examined the 12 that continuously measured and reported soil moisture. We found that experimental warming reduced soil moisture across all warming types, with substantial variation among experiments (Figure 5, Table S15). The drying effect was greatest in studies using both forced air and soil warming (in which this reduction occurred at a rate of 0.36% per degree of target warming,Table S14), compared with those using infrared heating (in which this reduction occurred at a rate of 0.36% per degree of target warming,Table S14). Soil moisture can be difficult to measure, with dramatic variation in space and time (Famiglietti *et al.*, 1999; Teuling & Troch, 2005), but these results suggest that soil moisture is unavoidably affected by changing temperatures, even when active warming experiments may not be explicitly designed to manipulate soil moisture.

Warming and precipitation treatments, and their secondary effects on soil moisture and other abiotic factors, can also alter the biotic environment, which may produce cascading effects. Many studies have found shifts from herbaceous to woody plant communities over time with experimental warming (e.g., Rollinson & Kaye, 2012; McDaniel *et al.*, 2014b,a; Harte *et al.*, 2015); this, in turn, can alter microbial and herbaceous plant communities. These community shifts may change competitive dynamics and affect resource levels, such as moisture, carbon, and nutrient levels in the soil (McDaniel *et al.*, 2014b,a; Harte *et al.*, 2015), and may feedback to affect microclimate (Harte *et al.*, 2015).

The presence of these feedback effects is both a strength and a challenge of climate change experiments. They may represent important and ecologically realistic effects that might not have been apparent without the *in situ* field experiment. Alternatively, they may represent artifacts that are unlikely to occur outside of an experimental context. Quantifying, interpreting, and reporting these non-temperature effects in experiments is critical to distinguishing this and to understanding mechanisms underlying observed biological responses to climate change.

The widespread presence of secondary effects of climate manipulations highlights the importance of measuring environmental conditions at the plot level, and using these measurements in analysis and interpretation of results. Many climate change experiments—including 10 of the 15 in the MC3E database —analyze warming and/or precipitation treatments as simple categorical predictors (e.g., as in a two-way ANOVA, Table S1). Our findings, however, demonstrate a clear need for alternative modelling approaches to fully understand the experimental results and to make mechanistic links between changes in climate and ecological responses. One straightforward alternative is to include the continuous climate data (e.g., plot-level temperatures) as predictors of the focal response variable, such as phenological state or species density (e.g., Marchin *et al.*, 2015; Pelini *et al.*, 2014).

Ecological implications

We have highlighted a suite of factors that complicate interpretation of warming experiments. These largely non-target alterations, analogous to the "hidden treatments" described by Huston (1997) in biodiversity experiments, which can be seen as both a strength and a challenge of field experiments, are likely to have biological implications for many of the responses studied in warming experiments (e.g., Figure 5). Interpretation of experimental climate change effects on biological responses may be misleading because the intended climate treatments (i.e., categorical comparisons or target warming levels) are generally used as explanatory variables in analyses (Table S1). The interpretation is likely to be altered by using fine-scale, measured climate as explanatory variables. Detailed examination of multiple microclimate variables (e.g., plot-level temperature and soil moisture) will allow a more complete understanding of the indirect, as well as direct, effects of treatments on abiotic and biotic drivers of focal responses.

Biological responses may be muted (Figure 5b) or exaggerated (Figure 5c) in experiments when direct and indirect effects of climate manipulations interact. Plant phenology provides one example of a biological response that appears to be muted in experiments versus observational studies (Figure 5b). This is because phenology has a complex dependence on temperature and water availability (as well as other factors, Davis *et al.*, 2015). Although phenology is generally advanced by higher spring temperatures, it can also be delayed by increased winter temperatures (which delay endodormancy break). In addition, reduced water availability during the spring can slow cell elongation and delay budburst (Peñuelas *et al.*, 2004; Ourcival & Rambal, 2011; Craine *et al.*, 2012; Matthews & Mazer, 2016). Effects of these different drivers may be responsible for discrepancies between observational and experimental phenological responses to warming, which have been observed in diverse species, warming types, and locations (Wolkovich *et al.*, 2012).

We tested if temperature sensitivity from experiments varies when estimated using simple target warming compared to when measured plot-level climate variables are used. With data in the MC3E database, we fit two separate models, one using target warming, and one using measured climate. For both models, the response variable was budburst day of year and we accounted for non-independence due to site and year with random effects (see Supplemental Materials for details).

Temperature sensitivity estimates from the two modeling approaches varied three-fold. The target warming model estimated temperature sensitivity of budburst to be -2.01 days/°C (95% CI -2.17,-1.86; i.e., budburst shifts earlier by two days per °C of warming) (Table S15, solid black line in Figure 6), whereas the measured climate model estimated temperature sensitivity of budburst to be -6.22 days/°C (95% CI:-7.034,-5.41; Table S15). Further, all measured climate models with both temperature and moisture had improved model fit compared to the target warming model (Table S16). The best-fit model included mean daily minimum above-ground temperature, mean winter soil moisture, and their interaction as explanatory variables, suggesting that these variables and their interaction are important drivers of budburst timing (Tables S15, S16). In addition, the measured climate model estimated a significant negative effect of soil moisture on budburst of -1.35 days/% VWC (95% CI: -1.58,-1.13; Table S15, Figure 6). This negative effect is expected, if reducing moisture delays budburst (Table S15, Figure 6), and is consistent with previous work showing that budburst requires water uptake (Essiamah & Eschrich, 1986).

The increase in estimated temperature sensitivity with measured (rather than target) temperature has two major causes. First, target warming generally exceeds the measured above-ground temperature differences between treatment and control plots (Figure 2). Second, experimental warming dries out the soil in addition to increasing temperatures, and both climate variables affect the timing of budburst. Decreasing soil moisture has a delaying effect on budburst phenology, opposing the advancing effect of rising temperatures (Figure 5b). This example shows how the common method of using target warming alone to understand biological responses is likely to yield inaccurate estimates of temperature sensitivity in warming experiments. In this case, the underestimation may be substantial enough to account for the previously observed discrepancy between observational and experimental phenological responses to warming, though further investigation is required, for example across additional phenophases (Wolkovich *et al.*, 2012).

Accounting for both direct and indirect effects of warming is critical for accurate interpretation of the consequences of climate change (Kharouba *et al.*, 2015). A critical question is the extent to which abiotic and biotic effects are accurate forecasts of future shifts that are likely to occur with climate change, or due to artifacts that are unlikely to occur outside of experimental systems (Moise & Henry, 2010; Diamond *et al.*, 2013). For many important climatic and ecological metrics, experimental findings of abiotic and biotic effects appear to be consistent with observations. Altered above-ground daily temperature range (i.e. temperature minima changing more than maxima, Table S3) with experimental warming is consistent with observed changes in many places, at least for some time periods. Minimum temperatures increased more rapidly than maximum temperatures, reducing above-ground daily temperature range strongly and significantly from 1950-1980 (though the trends have been largely insignificant from 1980 onward Thorne *et al.*, 2016; Vose *et al.*, 2005). In addition, shifts from non-woody to woody vegetation, coupled with declines in soil carbon, are two effects of warming, observed in both experimentally warmed plots over the short-term and ambient controls over decades of climate warming at a sub-alpine meadow site (Harte *et al.*, 2015). The acclimation response of leaf respiration to temperature (Aspinwall *et al.*, 2016; Reich *et al.*, 2016), and responses of soil respiration to warming (Carey *et al.*, 2016), also appear to be consistent across experiments and observations. These cases suggest that many responses observed in climate change experiments are likely to be accurate harbingers of future biological responses to climate change, with the caveat that short-term responses frequently differ from long-term responses (Andresen *et al.*, 2016).

In other cases, however, some of the non-temperature effects observed in climate change experiments may be potential experimental artifacts. For example, soil drying in conjunction with future warming is forecasted in some regions, such as the southwestern United States, mainly because of reductions in precipitation and increased evaporative demand with warmer air (Dai, 2013; Seager *et al.*, 2013). The northeastern United States, on the other hand, has been trending wetter over time (Shuman & Burrell, 2017), even though temperatures have warmed. Future changes in soil moisture are certain, and likely to vary by region, season, and even soil depth (Seager *et al.*, 2014; Berg *et al.*, 2017). Thus, it is not safe to assume that the soil drying observed in warming experiments is necessarily likely to occur with future warming; rather, this non-temperature effect of experimental warming deserves explicit analysis and interpretation. The altered light, wind, and herbivory patterns documented under experimental infrastructure (Kennedy, 1995; Moise & Henry, 2010; Wolkovich *et al.*, 2012; Hoeppner & Dukes, 2012; Clark *et al.*, 2014b) represent other non-temperature effects that may be potential experimental artifacts and are worth quantifying in future analyses to provide improved estimates of temperature sensitivity.

Conclusions

As climate change continues across the globe, ecologists are challenged to not only document impacts but to make quantitative, robust predictions. Our ability to meet this challenge requires a nuanced mechanistic understanding of how climate directly and indirectly alters biological processes. Climate change experiments, which have been underway for nearly four decades (e.g., Tamaki *et al.*, 1981; Carlson & Bazzaz, 1982; Melillo *et al.*, 2017), provide invaluable information about biological responses to climate change. Yet the full range of changes in environmental conditions imposed by these experiments is rarely presented. We have compiled the first database of fine-scale climate data from multiple warming experiments and shown how time, space, experimental artifacts, and secondary effects of treatments may hinder simple interpretations of these experimental results. We hope this work provides a foundation for gaining the most knowledge and utility from existing experiments via robust analyses, for designing better experiments and models in the future (see Box 1), and for improved understanding of biological responses and feedbacks in a changing world.

Acknowledgements

We are grateful to those who shared their experimental climate data with us, allowing it to be included in the MC3E database. We thank the Radcliffe Institute for Advanced Study at Harvard University, which provided funding for an Exploratory Seminar at which the ideas in this paper were conceived. This research was also supported by the National Science Foundation (NSF DBI 14-01854 to A.K.E.). Any opinion, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the National Science Foundation.

Box 1: Recommendations for future climate change experiments

1. *Collect and analyze fine-scale climate data.* This includes analyzing and interpreting minimum and maximum values, as well as variance and critical thresholds (e.g., the number and duration of freeze-thaw events and accumulated chilling hours, McDaniel *et al.*, 2014b; Vasseur *et al.*, 2014). We suggest saving the raw data from data loggers (often collected at hourly or higher resolution) to allow quantification of variance (and other summaries) at different temporal resolutions. In assessing which frequency of measurements is most appropriate for analyses (e.g., hourly, twice daily), it is critical to consider the chronobiology of the event and organisms of interest. For ants, this might mean that temperatures be monitored every minute (Helm & Shavit, 2017); for bacteria, even more frequently.

2. *Analyze measured climate variables rather than targets.* There can be substantial variation in the effects of warming and precipitation treatments among plots and across time (Figure 2). Analyzing measured climate will allow much more in-depth understanding of the drivers and biological effects of variation in temperature and moisture.

3. *Publish high quality, usable data and metadata.* Given that climate manipulations are logistically challenging and expensive (Aronson & McNulty, 2009), and that they often produce a large volume of fine-scale climate data, good curation and data sharing will ensure wider use and deeper understanding of these valuable data. When studying biological implications of a global challenge as large as climate change, progress will come from designing and reporting experiments in ways that facilitate an eventual global data set.

4. *Include both structural and ambient controls* and collect, use, and report climate and biological data within them. Fewer than half of the studies in our MC3E database reported climate data from these two control types (6 out of 15 studies); however, all experiments that did include both control types showed significant effects of infrastructure (Figure 3).

5. *Design relevant manipulations* by consulting observational records and forecasts, including seasonal and annual variation in projected warming. When it is not possible or desirable to match anticipated changes in climate, studies should report how imposed treatments compare to projected changes and past observations (e.g., Hoover *et al.*, 2014; Zhu *et al.*, 2016). In addition, if continuous treatments are not applied throughout the study, the seasonality and timing of treatments should be explicitly reported and the climate should be monitored throughout.

6. *Maximize the duration of climate change experiments* by running some experiments for as long as possible, since the magnitude of climate change treatments can vary considerably among years (Figure 2. In addition, long-term responses of individuals and populations can differ from transient responses (Saleska *et al.*, 2002; Franklin, 1989; Giasson *et al.*, 2013; Harte *et al.*, 2015). We were only able to acquire data extending for 5 years or more for one study in the MC3E database (exp01), restricting our ability to investigate the effect of study length on experimental climate change. Well-designed and well-supported longer warming experiments will allow investigation of how inter-annual variations interact with climate change treatments, particularly when combined with observational studies and modeling (Luo *et al.*, 2011).

7. *Conduct additional syntheses across studies.* As more detailed data become published from experimental climate change studies in diverse ecosystems, meta-analyses of these data will allow further understanding of the ways that microclimate and biotic interactions are affected by active warming. For example, it would be useful to compare microclimates in studies using infrared warming applied with constant voltage versus infrared warming that varies voltage based on measured temperatures.

References

Andresen, L.C., Müller, C., de Dato, G., Dukes, J.S., Emmett, B.A., Estiarte, M., Jentsch, A., Kröel-Dulay, G., Lüscher, A., Niu, S. *et al.* (2016). Shifting impacts of climate change: long-term patterns of plant response to elevated co2, drought, and warming across ecosystems. In: *Advances in ecological research*. Elsevier, vol. 55, pp. 437–473.

Aronson, E.L. & McNulty, S.G. (2009). Appropriate experimental ecosystem warming methods by ecosystem, objective, and practicality. *Agricultural and Forest Meteorology*, 149, 1791–1799.

Aspinwall, M.J., Drake, J.E., Campany, C., Vårhammar, A., Ghannoum, O., Tissue, D.T., Reich, P.B. & Tjoelker, M.G. (2016). Convergent acclimation of leaf photosynthesis and respiration to prevailing ambient temperatures under current and warmer climates in eucalyptus tereticornis. *New Phytologist*, 212, 354–367.

Berg, A., Sheffield, J. & Milly, P.C. (2017). Divergent surface and total soil moisture projections under global warming. *Geophysical Research Letters*, 44, 236–244.

Bokhorst, S., Huiskes, A., Aerts, R., Convey, P., Cooper, E.J., Dalen, L., Erschbamer, B., Gudmundsson, J., Hofgaard, A., Hollister, R.D. *et al.* (2013). Variable temperature effects of open top chambers at polar and alpine sites explained by irradiance and snow depth. *Global Change Biology*, 19, 64–74.

Carey, J.C., Tang, J., Templer, P.H., Kroeger, K.D., Crowther, T.W., Burton, A.J., Dukes, J.S., Emmett, B., Frey, S.D., Heskel, M.A. *et al.* (2016). Temperature response of soil respiration largely unaltered with experimental warming. *Proceedings of the National Academy of Sciences*, 113, 13797–13802.

Carlson, R.W. & Bazzaz, F.A. (1982). Photosynthetic and growth response to fumigation with so2 at elevated co2 for c3 and c4 plants. *Oecologia*, 54, 50–54.

Chuine, I., Bonhomme, M., Legave, J.M., García de Cortázar-Atauri, I., Charrier, G., Lacointe, A. & Améglio, T. (2016). Can phenological models predict tree phenology accurately in the future? the unrevealed hurdle of endodormancy break. *Global Change Biology*, 22, 3444–3460.

Chuine, I., Morin, X., Sonié, L., Collin, C., Fabreguettes, J., Degueldre, D., Salager, J.L. & Roy, J. (2012). Climate change might increase the invasion potential of the alien c4 grass setaria parviflora (poaceae) in the mediterranean basin. *Diversity and Distributions*, 18, 661–672.

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

Clark, J.S., Salk, C., Melillo, J. & Mohan, J. (2014b). Tree phenology responses to winter chilling, spring warming, at north and south range limits. *Functional Ecology*, 28, 1344–1355.

Cleland, E.E., Chiariello, N.R., Loarie, S.R., Mooney, H.A. & Field, C.B. (2006). Diverse responses of phenology to global changes in a grassland ecosystem. *Proceedings of the National Academy of Sciences of the United States of America*, 103, 13740–13744.

Cox, P.M., Betts, R.A., Jones, C.D., Spall, S.A. & Totterdell, I.J. (2000). Acceleration of global warming due to carbon-cycle feedbacks in a coupled climate model. *Nature*, 408, 184–187.

Craine, J.M., Wolkovich, E.M., Towne, E.G. & Kembel, S.W. (2012). Flowering phenology as a functional trait in a tallgrass prairie. *New Phytologist*, 193, 673–682.

Dai, A. (2013). Increasing drought under global warming in observations and models. *Nature Climate Change*, 3, 52–58.

Davis, C.C., Willis, C.G., Connolly, B., Kelly, C. & Ellison, A.M. (2015). Herbarium records are reliable sources of phenological change driven by climate and provide novel insights into species' phenological cueing mechanisms. *American journal of botany*, 102, 1599–1609.

De Boeck, H.J., Vicca, S., Roy, J., Nijs, I., Milcu, A., Kreyling, J., Jentsch, A., Chabbi, A., Campioli, M., Callaghan, T. *et al.* (2015). Global change experiments: challenges and opportunities. *BioScience*, 65, 922–931.

Diamond, S.E., Penick, C.A., Pelini, S.L., Ellison, A.M., Gotelli, N.J., Sanders, N.J. & Dunn, R.R. (2013). Using physiology to predict the responses of ants to climatic warming. *Integrative and comparative biology*, 53, 965–974.

Dukes, J.S. & Mooney, H.A. (1999). Does global change increase the success of biological invaders? *Trends in Ecology and Evolution*, 14, 135–139.

Dunne, J.A., Harte, J. & Taylor, K.J. (2003). Subalpine meadow flowering phenology responses to climate change: integrating experimental and gradient methods. *Ecological Monographs*, 73, 69–86.

Essiamah, S. & Eschrich, W. (1986). Water uptake in deciduous trees during winter and the role of conducting tissues in spring reactivation. *IAWA Journal*, 7, 31–38.

Ettinger, A. & Wolkovich, E. (2018). Microclimate from climate change experiments (MC3E). doi:10.5063/F1QV3JQR.

Famiglietti, J., Devereaux, J., Laymon, C., Tsegaye, T., Houser, P., Jackson, T., Graham, S., Rodell, M. & Oevelen, P.V. (1999). Ground-based investigation of soil moisture variability within remote sensing footprints during the southern great plains 1997 (sgp97) hydrology experiment. *Water Resources Research*, 35, 1839–1851.

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

Field, C.B., Lobell, D.B., Peters, H.A. & Chiariello, N.R. (2007). Feedbacks of terrestrial ecosystems to climate change. *Annual Review of Environment and Resources*, 32, 1–29.

Franklin, J.F. (1989). Importance and justification of long-term studies in ecology. In: *Long-term studies in ecology*. Springer, pp. 3–19.

Friedlingstein, P., Meinshausen, M., Arora, V.K., Jones, C.D., Anav, A., Liddicoat, S.K. & Knutti, R. (2014). Uncertainties in cmip5 climate projections due to carbon cycle feedbacks. *Journal of Climate*, 27, 511–526.

Giasson, M.A., Ellison, A.M., Bowden, R., Crill, P.M., Davidson, E., Drake, J., Frey, S., Hadley, J., Lavine, M., Melillo, J. *et al.* (2013). Soil respiration in a northeastern us temperate forest: a 22-year synthesis. *Ecosphere*, 4, 1–28.

Gruner, D.S., Bracken, M.E., Berger, S.A., Eriksson, B.K., Gamfeldt, L., Matthiessen, B., Moorthi, S., Sommer, U. & Hillebrand, H. (2017). Effects of experimental warming on biodiversity depend on ecosystem type and local species composition. *Oikos*, 126, 8–17.

Hagedorn, F., Martin, M., Rixen, C., Rusch, S., Bebi, P., Zürcher, A., Siegwolf, R.T., Wipf, S., Escape, C., Roy, J. *et al.* (2010). Short-term responses of ecosystem carbon fluxes to experimental soil warming at the swiss alpine treeline. *Biogeochemistry*, 97, 7–19.

Hampe, A. (2004). Bioclimate envelope models: what they detect and what they hide. *Global Ecology and Biogeography*, 13, 469–471.

Harte, J., Saleska, S.R. & Levy, C. (2015). Convergent ecosystem responses to 23-year ambient and manipulated warming link advancing snowmelt and shrub encroachment to transient and long-term climate–soil carbon feedback. *Global Change Biology*, 21, 2349–2356.

Harte, J., Torn, M.S., Chang, F.R., Feifarek, B., Kinzig, A.P., Shaw, R. & Shen, K. (1995). Global warming and soil microclimate: Results from a meadow-warming experiment. *Ecological Applications*, 5, 132–150.

Helm, B. & Shavit, A. (2017). *Dissecting and reconstructing time and space for replicable biological research*, New Haven, CT: Yale University Press, pp. 233–249.

Hobbie, S.E., Shevtsova, A. & Chapin III, F.S. (1999). Plant responses to species removal and experimental warming in alaskan tussock tundra. *Oikos*, 84, 417–434.

Hoeppner, S.S. & Dukes, J.S. (2012). Interactive responses of old-field plant growth and composition to warming and precipitation. *Global Change Biology*, 18, 1754–1768.

Hoover, D.L., Knapp, A.K. & Smith, M.D. (2014). Resistance and resilience of a grassland ecosystem to climate extremes. *Ecology*, 95, 2646–2656.

Huston, M.A. (1997). Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. *Oecologia*, 110, 449–460.

Ibanez, I., Clark, J.S., Dietze, M.C., Feeley, K., Hersh, M., LaDeau, S., McBride, A., Welch, N.E. & Wolosin, M.S. (2006). Predicting biodiversity change: Outside the climate envelope, beyond the species-area curve. *Ecology*, 87, 1896–1906.

Johnson, P.D. & Besselsen, D.G. (2002). Practical aspects of experimental design in animal research. *ILAR Journal*, 43, 202–206.

Kearney, M. & Porter, W. (2009). Mechanistic niche modelling: combining physiological and spatial data to predict species' ranges. *Ecology Letters*, 12, 334–350.

Kennedy, A. (1995). Temperature effects of passive greenhouse apparatus in high-latitude climate change experiments. *Functional Ecology*, 9, 340–350.

Kharouba, H.M., Vellend, M., Sarfraz, R.M. & Myers, J.H. (2015). The effects of experimental warming on the timing of a plant–insect herbivore interaction. *Journal of Animal Ecology*, 84, 785–796.

Kimball, B. (2005). Theory and performance of an infrared heater for ecosystem warming. *Global Change Biology*, 11, 2041–2056.

Kimball, B.A., Conley, M.M., Wang, S., Lin, X., Luo, C., Morgan, J. & Smith, D. (2008). Infrared heater arrays for warming ecosystem field plots. *Global Change Biology*, 14, 309–320.

Luo, Y., Melillo, J., Niu, S., Beier, C., Clark, J.S., Classen, A.T., Davidson, E., Dukes, J.S., Evans, R., Field, C.B. *et al.* (2011). Coordinated approaches to quantify long-term ecosystem dynamics in response to global change. *Global Change Biology*, 17, 843–854.

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

Marion, G., Henry, G., Freckman, D., Johnstone, J., Jones, G., Jones, M., Levesque, E., Molau, U., Mølgaard, P., Parsons, A. *et al.* (1997). Open-top designs for manipulating field temperature in high-latitude ecosystems. *Global Change Biology*, 3, 20–32.

Matthews, E.R. & Mazer, S.J. (2016). Historical changes in flowering phenology are governed by temperature× precipitation interactions in a widespread perennial herb in western North America. *New Phytologist*, 210, 157–167.

McDaniel, M., Kaye, J. & Kaye, M. (2014a). Do "hot moments" become hotter under climate change? soil nitrogen dynamics from a climate manipulation experiment in a post-harvest forest. *Biogeochemistry*, 121, 339–354.

McDaniel, M., Wagner, R., Rollinson, C., Kimball, B., Kaye, M. & Kaye, J. (2014b). Microclimate and ecological threshold responses in a warming and wetting experiment following whole tree harvest. *Theoretical and Applied Climatology*, 116, 287–299.

Melillo, J., Frey, S., DeAngelis, K., Werner, W., Bernard, M., Bowles, F., Pold, G., Knorr, M. & Grandy, A. (2017). Long-term pattern and magnitude of soil carbon feedback to the climate system in a warming world. *Science*, 358, 101–105.

Menke, S.B., Harte, J. & Dunn, R.R. (2014). Changes in ant community composition caused by 20 years of experimental warming vs. 13 years of natural climate shift. *Ecosphere*, 5, 1–17.

Milcu, A., Puga-Freitas, R., Ellison, A.M., Blouin, M., Scheu, S., Girin, T., Frechet, G., Rose, L., Scherer-Lorenzen, M., Barot, S. *et al.* (2016). Systematic variability enhances the reproducibility of an ecological study. *bioRxiv(beta)*, p. 080119.

Moise, E.R. & Henry, H.A. (2010). Like moths to a street lamp: exaggerated animal densities in plot-level global change field experiments. *Oikos*, 119, 791–795.

Moorcroft, P.R. (2006). How close are we to a predictive science of the biosphere? *Trends in Ecology & Evolution*, 21, 400–407.

Morin, X., Roy, J., Sonié, L. & Chuine, I. (2010). Changes in leaf phenology of three european oak species in response to experimental climate change. *New Phytologist*, 186, 900–910.

Ohlemüller, R., Gritti, E.S., Sykes, M.T. & Thomas, C.D. (2006). Towards european climate risk surfaces: the extent and distribution of analogous and non-analogous climates 1931–2100. *Global Ecology and Biogeography*, 15, 395–405.

Ourcival, J. & Rambal, S. (2011). Phenological responses to extreme droughts in a Mediterranean forest. *Global Change Biology*, 17, 1036–1048.

Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of Ecology Evolution and Systematics*, 37, 637–669.

Pearson, R.G. & Dawson, T.P. (2004). Bioclimate envelope models: what they detect and what they hide - response to hampe (2004). *Global Ecology and Biogeography*, 13, 471–473.

Pelini, S., Diamond, S., Nichols, L., Stuble, K., Ellison, A.M., Sanders, N., Dunn, R. & Gotelli, N. (2014). Geographic differences in effects of experimental warming on ant species diversity and community composition. *Ecosphere*, 5, 1–12.

Pelini, S.L., Bowles, F.P., Ellison, A.M., Gotelli, N.J., Sanders, N.J. & Dunn, R.R. (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.

Peñuelas, J., Filella, I., Zhang, X., Llorens, L., Ogaya, R., Lloret, F., Comas, P., Estiarte, M. & Terradas, J. (2004). Complex spatiotemporal phenological shifts as a response to rainfall changes. *New Phytologist*, 161, 837–846.

Peterjohn, W.T., Melillo, J.M., Bowles, F.P. & Steudler, P.A. (1993). Soil warming and trace gas fluxes: experimental design and preliminary flux results. *Oecologia*, 93, 18–24.

Price, M.V. & Waser, N.M. (1998). Effects of experimental warming on plant reproductive phenology in a subalpine meadow. *Ecology*, 79, 1261–1271.

Quinn, G.P. & Keough, M.J. (2002). *Experimental design and data analysis for biologists*. Cambridge University Press.

Reich, P.B., Sendall, K.M., Rice, K., Rich, R.L., Stefanski, A., Hobbie, S.E. & Montgomery, R.A. (2015). Geographic range predicts photosynthetic and growth response to warming in co-occurring tree species. *Nature Climate Change*, 5, 148–152.

Reich, P.B., Sendall, K.M., Stefanski, A., Wei, X., Rich, R.L. & Montgomery, R.A. (2016). Boreal and temperate trees show strong acclimation of respiration to warming. *Nature*, 531, 633–636.

Rich, R.L., Stefanski, A., Montgomery, R.A., Hobbie, S.E., Kimball, B.A. & Reich, P.B. (2015). Design and performance of combined infrared canopy and belowground warming in the b4warmed (boreal forest warming at an ecotone in danger) experiment. *Global change biology*, 21, 2334–2348.

Rollinson, C.R. & Kaye, M.W. (2012). Experimental warming alters spring phenology of certain plant functional groups in an early successional forest community. *Global Change Biology*, 18, 1108–1116.

Rollinson, C.R. & Kaye, M.W. (2015). Modeling monthly temperature in mountainous ecoregions: importance of spatial scale for ecological research. *Climate Research*, 64, 99–110.

Saleska, S.R., Shaw, M.R., Fischer, M.L., Dunne, J.A., Still, C.J., Holman, M.L. & Harte, J. (2002). Plant community composition mediates both large transient decline and predicted long-term recovery of soil carbon under climate warming. *Global Biogeochemical Cycles*, 16, 3–1–3–18.

Seager, R., Neelin, D., Simpson, I., Liu, H., Henderson, N., Shaw, T., Kushnir, Y., Ting, M. & Cook, B. (2014). Dynamical and thermodynamical causes of large-scale changes in the hydrological cycle over north america in response to global warming. *Journal of Climate*, 27, 7921–7948.

Seager, R., Ting, M., Li, C., Naik, N., Cook, B., Nakamura, J. & Liu, H. (2013). Projections of declining surface-water availability for the southwestern united states. *Nature Climate Change*, 3, 482.

Shaver, G.R., Canadell, J., Chapin, F.S., Gurevitch, J., Harte, J., Henry, G., Ineson, P., Jonasson, S., Melillo, J., Pitelka, L. *et al.* (2000). Global warming and terrestrial ecosystems: A conceptual framework for analysis ecosystem responses to global warming will be complex and varied. Ecosystem warming experiments hold great potential for providing insights on ways terrestrial ecosystems will respond to upcoming decades of climate change. Documentation of initial conditions provides the context for understanding and predicting ecosystem responses. *BioScience*, 50, 871–882.

Sheldon, K.S., Yang, S. & Tewksbury, J.J. (2011). Climate change and community disassembly: impacts of warming on tropical and temperate montane community structure. *Ecology Letters*, 14, 1191–1200.

Sherry, R.A., Zhou, X., Gu, S., 3rd, J.A.A., Schimel, D.S., Verburg, P.S., Wallace, L.L. & Luo, Y. (2007). Divergence of reproductive phenology under climate warming. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 198–202.

Shukla, J. & Mintz, Y. (1982). Influence of land-surface evapotranspiration on the earth's climate. *Science*, 215, 1498–1501.

Shuman, B.N. & Burrell, S.A. (2017). Centennial to millennial hydroclimatic fluctuations in the humid northeast United States during the holocene. *Quaternary Research*, 88, 1–11.

Spector, R. (2001). Progress in the search for ideal drugs. *Pharmacology*, 64, 1–7.

Stocker, T., Qin, D., Plattner, G., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, B. & Midgley, B. (2013). IPCC 2013. Climate change 2013: The physical science basis. Contribution of Working Group II to the fifth assessment report of the Intergovernmental Panel on Climate Change.

Swab, R.M., Regan, H.M., Keith, D.A., Regan, T.J. & Ooi, M.K.J. (2012). Niche models tell half the story: spatial context and life-history traits influence species responses to global change. *Journal of Biogeography*, 39, 1266–1277.

Tamaki, G., Weiss, M.A. & Long, G.E. (1981). Evaluation of plant density and temperature in predator-prey interactions in field cages. *Environmental Entomology*, 10, 716–720.

Templer, P.H., Phillips, N.G., Ellison, A.M. & Pelini, S.L. (2016). Ecosystem warming increases sap flow rates of northern red oak trees. *Ecosphere*, 7.

Teuling, A.J. & Troch, P.A. (2005). Improved understanding of soil moisture variability dynamics. *Geophysical Research Letters*, 32.

Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus, B.F.N., de Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S., Midgley, G.F., Miles, L., Ortega-Huerta, M.A., Peterson, A.T., Phillips, O.L. & Williams, S.E. (2004). Extinction risk from climate change. *Nature*, 427, 145–148.

Thorne, P., Donat, M., Dunn, R., Williams, C., Alexander, L., Caesar, J., Durre, I., Harris, I., Hausfather, Z., Jones, P. *et al.* (2016). Reassessing changes in diurnal temperature range: Intercomparison and evaluation of existing global data set estimates. *Journal of Geophysical Research: Atmospheres*, 121, 5138–5158.

Thuiller, W. (2004). Patterns and uncertainties of species' range shifts under climate change. *Global Change Biology*, 10, 2020–2027.

Urban, M.C., Tewksbury, J.J. & Sheldon, K.S. (2012). On a collision course: competition and dispersal differences create no-analogue communities and cause extinctions during climate change. *Proceedings of the Royal Society B-Biological Sciences*, 279, 2072–2080.

Vasseur, D.A., DeLong, J.P., Gilbert, B., Greig, H.S., Harley, C.D., McCann, K.S., Savage, V., Tunney, T.D. & O'Connor, M.I. (2014). Increased temperature variation poses a greater risk to species than climate warming. *Proceedings of the Royal Society of London B: Biological Sciences*, 281, 20132612.

Vose, R.S., Easterling, D.R. & Gleason, B. (2005). Maximum and minimum temperature trends for the globe: An update through 2004. *Geophysical Research Letters*, 32, e01221.

Williams, J.W. & Jackson, S.T. (2007). Novel climates, no-analog communities, and ecological surprises. *Frontiers in Ecology and the Environment*, 5, 475–482.

Williams, J.W., Jackson, S.T. & Kutzbacht, J.E. (2007). Projected distributions of novel and disappearing climates by 2100 AD. *Proceedings of the National Academy of Sciences of the United States of America*, 104, 5738–5742.

Wolkovich, E.M., Cook, B.I., Allen, J.M., Crimmins, T.M., Betancourt, J.L., Travers, S.E., Pau, S., Regetz, J., Davies, T.J., Kraft, N.J.B., Ault, T.R., Bolmgren, K., Mazer, S.J., McCabe, G.J., McGill, B.J., Parmesan, C., Salamin, N., Schwartz, M.D. & Cleland, E.E. (2012). Warming experiments underpredict plant phenological responses to climate change. *Nature*, 485, 494–497.

Zhu, K., Chiariello, N.R., Tobeck, T., Fukami, T. & Field, C.B. (2016). Nonlinear, interacting responses to climate limit grassland production under global change. *Proceedings of the National Academy of Sciences*, 113, 10589–10594.

Figures

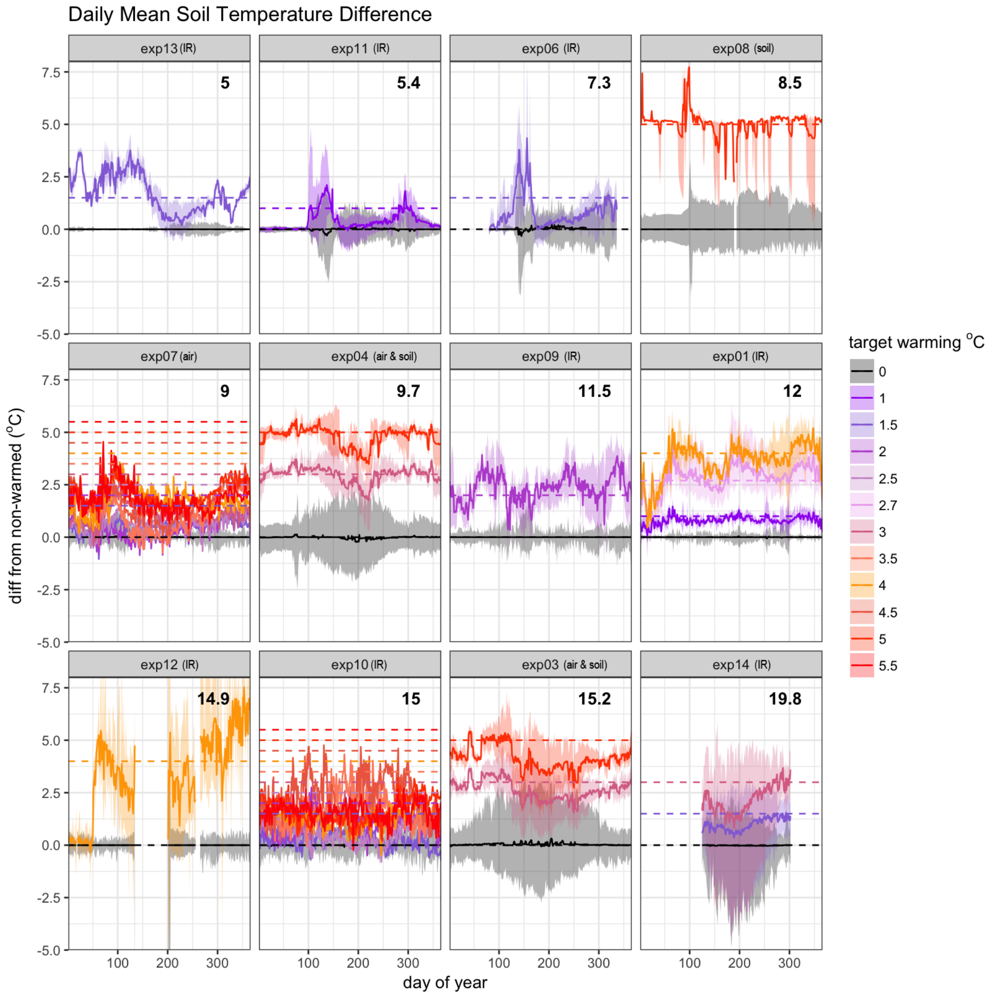

Figure 1: **Deviations in daily observed warming from mean control soil temperature for 10 study sites,** excluding data from plots that manipulated precipitation. We show soil, rather than above-ground, temperature, as this was the most frequently recorded temperature variable in the MC3E database. Solid lines show observed difference between warming treatment (colors) and control (black) plots, averaged across replicates and years; shading shows 95% confidence intervals. Dashed lines represent target warming levels. Two sites not shown here did not monitor soil temperature. Experimental sites are ordered by low to high mean annual soil temperature (shown in the upper right corner of each panel). The heating type is listed in parantheses next to the site number (IR= infrared, soil= soil cables, air= forced air).

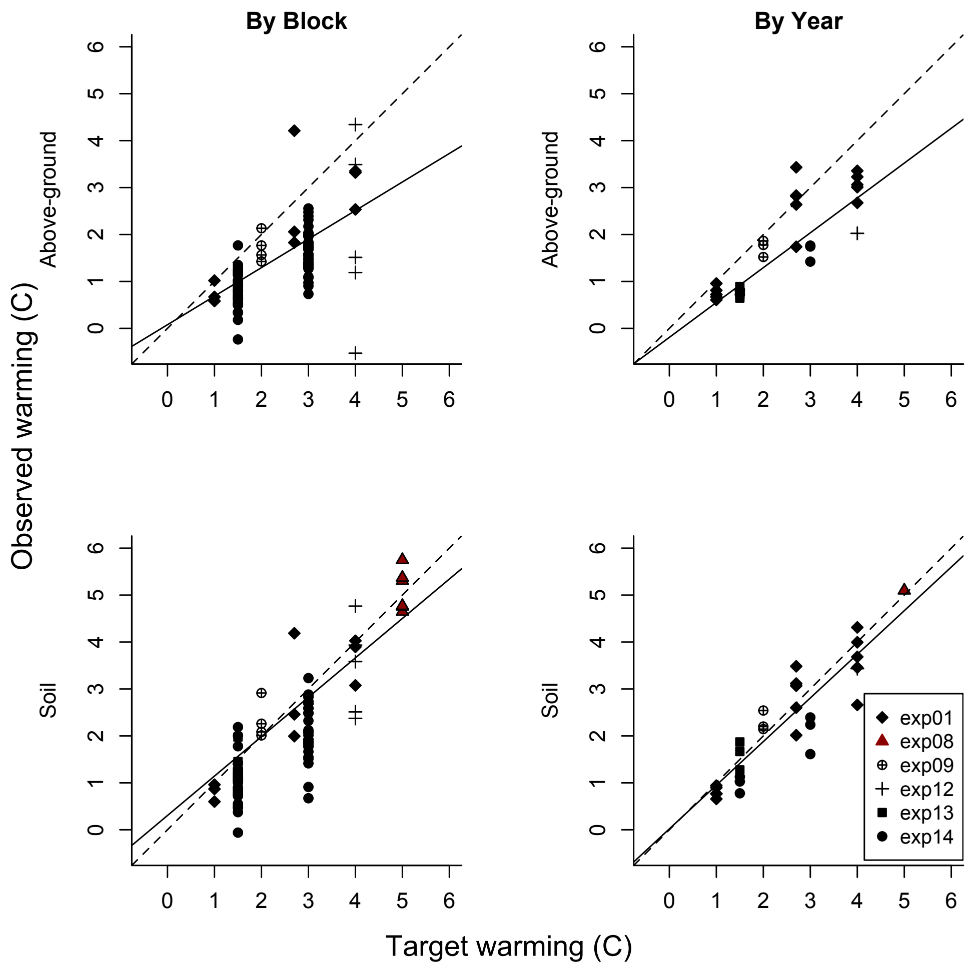

Figure 2: **Observed warming over space and time, for above-ground and soil temperatures,** excluding data from plots that manipulated precipitation. Above-ground temperature includes air, canopy, and surface temperature. Points represent the difference between treatment and control plots by block (i.e., one data point per block) and by year (i.e., one data point per year). 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). Colors vary by heating type: gray represents infrared; red represent soil warming cables. See Supplemental Materials (especially Tables S4 and S5) for details.

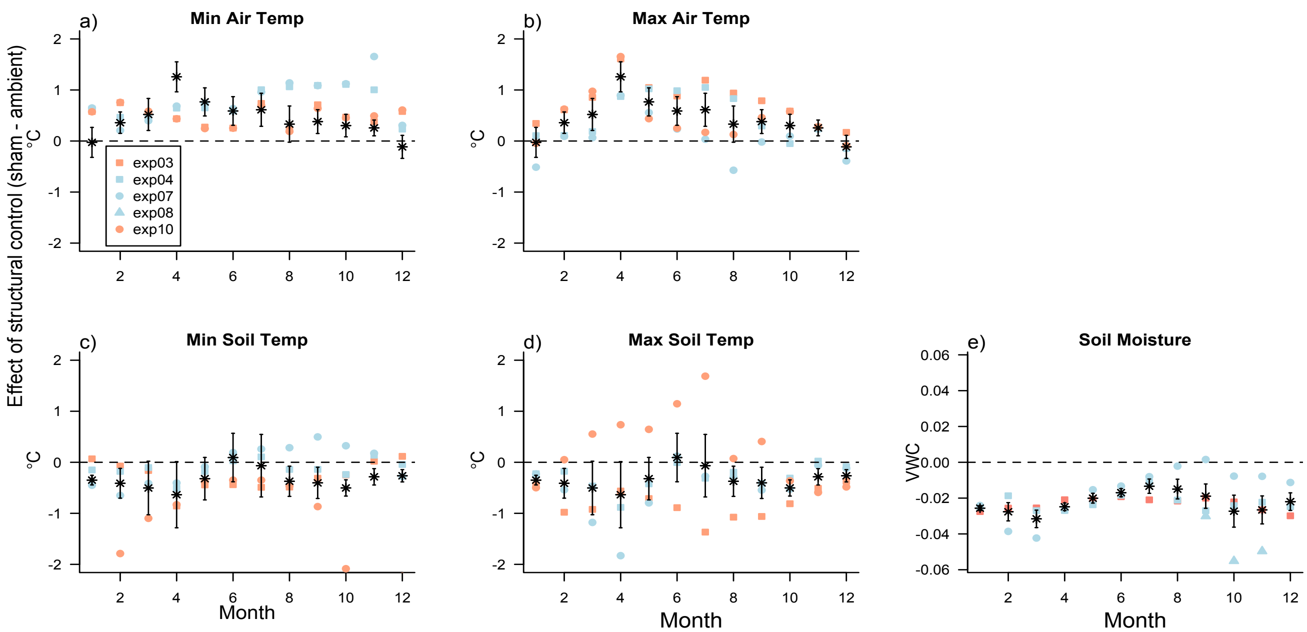

Figure 3: **Deviations in measured abiotic variables by month in structural controls compared to ambient controls** (i.e., with no control chambers or warming infrastructure in place). Above-ground temperatures were higher (a,b), whereas soil temperature (c,d) and soil moisture (e) were lower in structural controls compared with ambient controls. We show overall (fixed) effects in black from monthly mixed-effects models; site-level random effects are shown by symbols in blue (for the three studies conducted at Harvard Forest in Massachusetts, USA) and pink (the two studies conducted at Duke Forest in North Carolina, USA). Shapes vary by heating type: triangles represent soil warming cables, circles represent forced air; squares represent combined soil warming and forced air heating. See Supplemental Materials for details (Tables S6-S11).

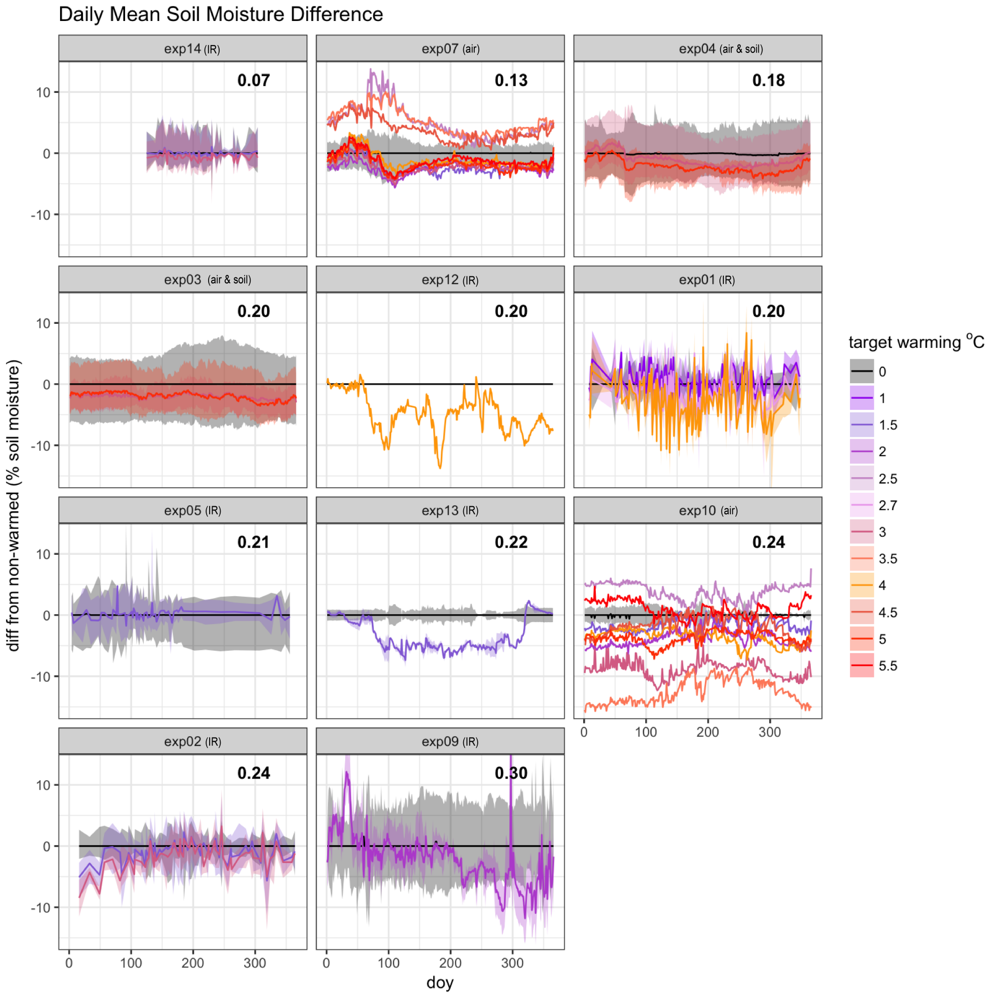

Figure 4: **Deviations in daily observed soil moisture,** shown for the nine study sites that continuously monitored soil moisture, excluding data from plots that manipulated precipitation. Black lines represent control plots, and colored lines represent warming treatments with various target warming levels. The number of temperature treatment levels vary from one (e.g., exp08, exp11) to nine (exp07 and exp10, which used an unreplicated regression design). Experimental sites are ordered by low to high mean annual soil moisture (shown in the upper right corner of each plot). All experiments measured soil moisture in volumetric water content, as a percentage of the soil volume in the sample, scaled from 0 to 100.

.

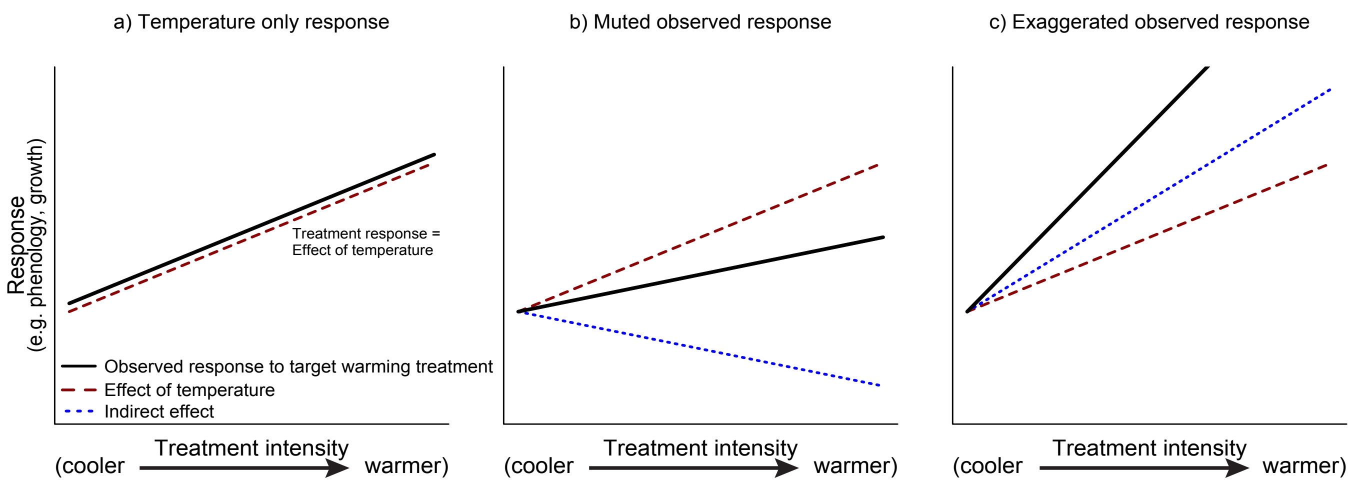

Figure 5: **Theoretical biological responses to experimental warming and their interpretation**. Direct responses to temperature alone (a) can be easily understood. Complications arise when biological responses are a mix of the direct temperature and indirect nontemperature effects of experimental warming. Then experimental warming may cause biological responses to be muted (b) or exaggerated (c). Quantifying, interpreting, and reporting these non-temperature effects in experiments is critical, and their presence is both a strength and a challenge of climate change experiments. They may represent ecologically realistic effects that might not have been predicted without the *in situ* field experiment. Alternatively, they may represent artifacts that are unlikely to occur outside of an experimental context. Slopes of these example lines assume a linear response with additive direct and indirect effects. The relationship between these effects could be more complex (e.g., nonlinear; antagonistic, multiplicative, or otherwise interactive).

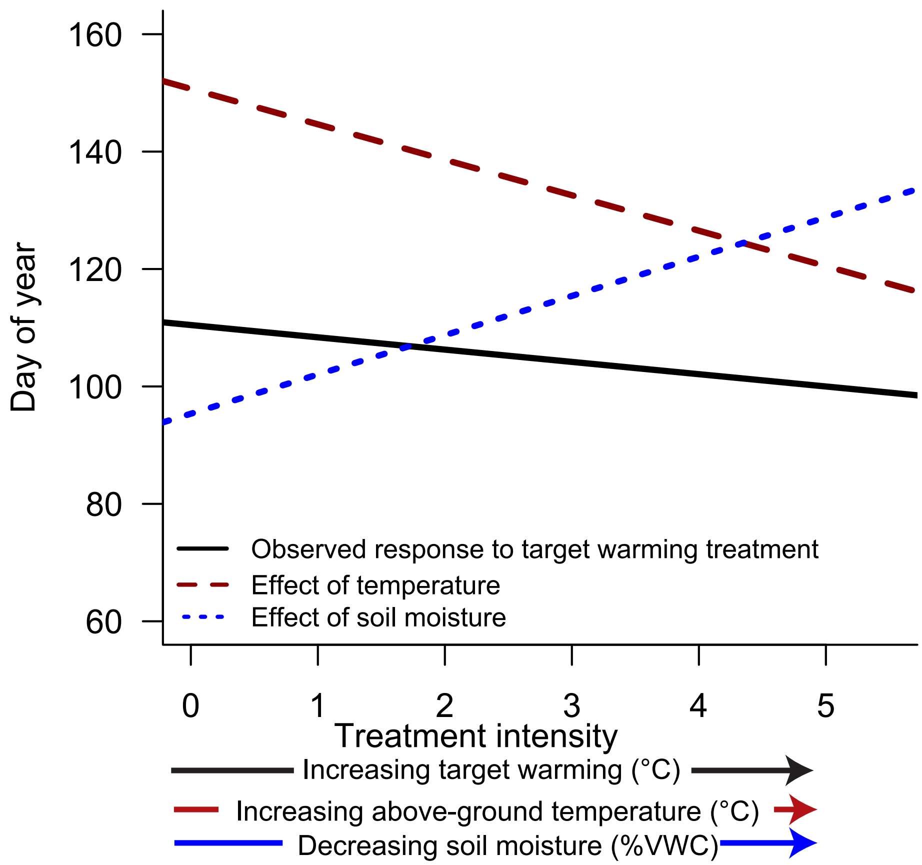

Figure 6: **Observed response of budburst day of year to experimental climate change** is an example of a muted response: the observed response to increasing treatment intensity (i.e., the coefficient of a model fit with only target temperature as the explanatory variable, black line; units for x-axis are °C of target warming) suggests a weaker temperature sensitivity than the effect of temperature in a more biologically accurate (and better-fitting) model that includes both measured above-ground temperature (dashed red line, for which x-axis units are °C of measured temperature) and soil moisture (dotted blue line, for which x-axis units are % VWC, decreasing from left to right in conjunction with warming intensity), as well as their interaction. This is because experimental warming dries out the soil in addition to increasing temperatures, and both climate variables affect the timing of budburst. Whereas increasing temperatures advance budburst, decreasing soil moisture has a delaying effect. A critical question is whether the soil drying that occurs in warming experiments is consistent with forecasts with climate change, since soil moisture trends are expected to vary by reegion, season, and soil depth (Berg *et al.*, 2017). Analysis includes all studies that monitored budburst, and measured soil moisture and above-ground temperature (exp01,exp03,exp04,exp07,exp10); see Supplemental Materials, especially Tables S14 & S15, for additional details.