How do climate change experiments actually change 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

January 18, 2017

Introduction

Ongoing climatic changes are causing dramatic alterations to Earth's biota, increasingly altering the physiology, distribution, and abundance of organisms, and resulting in cascading community and ecosystem 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 remains, however, about how particular individuals, populations, species, communities, and ecosystems will respond as shifts in temperature and precipitation regimes become more extreme. Predicting these biological responses to current and future climate change, as well as how they will feedback to affect earth's climate and ecosystem services, are among the most significant challenges facing biologists today.

Researchers have sought to understand and forecast biological responses through a variety of strategies, including observational studies, model-based approaches, and experiments. Much work has focused on temperature because increased greenhouse gas emissions have a relatively straightforward effect on this climatic variable, at least as compared to precipitation (Stocker *et al.*, 2013). Observational studies typically correlate observed biological patterns with trends in climate; however, it is challenging to disentangle the causal effect of warming from other factors that have also changed, such as successional stage or land use. In addition, future climate change is expected to result in conditions that are outside those that have been observed historically (Ohlemüller *et al.*, 2006; Williams & Jackson, 2007; Williams *et al.*, 2007; Stocker *et al.*, 2013). Modelling techniques can be useful in this regard, but these methods often rely on untested assumptions, for example, that climate limits performance (Pearson & Dawson, 2004; Ibanez *et al.*, 2006; Swab *et al.*, 2012). Experiments are therefore a critical component of the biologist's toolkit for understanding climate change impacts, and are often considered the "gold standard" of knowledge (e.g. Box *et al.*, 1978; Gelman, 2014). In situ climate warming experiments allow effects of temperature and precipitation to be isolated from other environmental changes that can confound conclusions drawn from observational data sets. In addition, a range of warming and precipitation treatments can be applied, such that potential non-linear responses may be evaluated. Compared with controlled growth-chamber experiments, field-based experiments offer the possibility of preserving important, but unknown or unquantified, in situ biotic and abiotic drivers and interactions. Climate change experiments in the field may therefore be able to elucidate many biological responses to future climate change.

Experimental climate manipulations take a variety of forms, manipulating temperature, precipitation, atmospheric CO2, and other variables using a wide range of techniques (Shaver *et al.*, 2000; Aronson & McNulty, 2009). In the field, temperature increases can be simulated using "passive" warming infrastructure, such as open-top chambers, that trap energy already available in the environment, or "active" warming methods, which heat ecosystems using external energy inputs (e.g. gas-powered forced air heaters, electrical-powered soil warming cables or infrared heaters Shaver *et al.*, 2000). Many field experiments have explored biological responses to interacting environmental changes: active warming methods have been combined with precip-

itation manipulations, (e.g. snow removal, water additions and reductions), in an effort to create climatic conditions such as those forecasted under future climate change scenarios (e.g. Price & Waser, 1998; Cleland *et al.*, 2006; Sherry *et al.*, 2007; Rollinson & Kaye, 2012).

Seeking to prepare for future, altered biological conditions, scientists and others often attempt to extrapolate the results of in situ climate change experiments to forecast how organisms and ecosystems will respond to particular climate change scenarios. Even in cases where extrapolation is not the explicit goal of climate change experiments, these studies are often used to draw conclusions about how anthropogenic warming will affect species' performance (e.g. growth and survival) and distributions(Dukes & Mooney, 1999; Hobbie *et al.*, 1999; Reich *et al.*, 2015; Gruner *et al.*, 2016). Despite these applications, a detailed assessment of how different types of experimental warming alter the environmental conditions experienced by organisms, and the extent to which these conditions accurately simulate current field conditions and anticipated climate change, is lacking. In addition comparing experimental results with observations and forecasts, there is a need to reconcile experimental results across the diverse methods, locations, and species to date.

A detailed understanding of how climate change experiments actually alter climate is critical in order to realize the forecasting potential of climate change experiments. Here, we use plot-level daily microclimate data from 12 climate change experiments that manipulate temperature and precipitation to demonstrate the direct and indirect ways in which environmental conditions are altered by active warming technologies. We then highlight the challenges associated with quantifying and interpreting biological responses to these climate manipulations, and using these interpretations to forecast more widespread responses to contemporary climate change. Finally, we use findings from our synthesis to make recommendations for future climate change experiments. We focus on in situ active warming manipulations, because recent analyses indicate that active warming methods are the most controlled, consistent, and "true to climate change predictions" (Kimball, 2005; Kimball *et al.*, 2008; Aronson & McNulty, 2009; Wolkovich *et al.*, 2012). The data we use were collected between 1991 and 2014 from North American and European climate change experiments (Figure 1) and have been merged into a new, publicly available Climate from Climate Change Experiments (C3E) database (see Supplemental Materials for details).

Complications in extrapolating 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 the experiment. However, biologists are generally interested primarily in the biological responses associated with each treatment (e.g. growth, abundance, or phenology of a species). Not surprisingly, then, 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 or not that mean change matched their target level of change (e.g. Price & Waser, 1998; Clark *et al.*, 2014a,b; Rollinson & Kaye, 2012).

Though the published focus is often on shifts in the mean, the imposed climate manipulations actually result in much more complex climatic shifts, for several reasons that we discuss in more detail below. First, in addition to shifting mean temperatures, experimental warming treatments may alter the temperature range and variance as well. Second, the magnitude of change in these manipulations is likely to vary in time and space. Third, the equipment required to conduct these manipulations can also alter climate at the plot level. All of these complications challenge our interpretation of how experimental warming studies can be applied to forecast effects of climate change, and we discuss them in more detail below.

Treatments alter the temperature range and variance, as well as the mean

Warming treatments affect minimum and maximum temperatures in different ways, and also alter the variance in temperature. For example, we found that active warming increased minimum air temperature by, on average, 0.84 degrees C per degree of warming target, whereas maximum temperature increased only 0.51 degrees, on average, per degree of target warming. We also found higher coefficients of variation in air temperatures in actively warmed plots, compared with controls at the same sites in the same years. This was true for both minimum and maximum air temperatures (Figure 2). These differences in temperature range and variation, appear to be small, but they may be important. For example, even slight changes in temperature can alter critical biological thresholds, such as freeze-thaw cycles (McDaniel *et al.*, 2014). Other studies have also documented that warming treatments may affect diurnal versus night temperatures differently (Shen *et al.*, 2016; Matthews & Mazer, 2016). (Question for everyone (Lizzie/Ben/Miriam/Ann Marie/Aaron/Yann/Isabelle/Jeff/Christy): this paragraph needs help! I have had trouble coming up with a good figure to go with this point, and have found relatively small but significant differences in temperature range and coefficient of variation. Alternatively, do you think this paragraph should be cut?)

Treatments vary over time

The common practice of reporting only the mean temperature difference, across the duration of the study, may also hide variations in daily, seasonal, and annual temperatures among treatments. For example, as described above, warming treatments can cause decreases in the diurnal temperature range within experimental plots, compared with ambient conditions (Hoeppner & Dukes, 2012). This may be similar to what is projected for parts of world; however, this will likely vary spatially, as some regions have experienced higher daytime warming than nighttime warming, whereas others have experienced the opposite (Stocker *et al.*, 2013).

In addition to daily fluctuations, there are frequently strong seasonal variations in experimental warming effects (Figure 3). (Christy: Could you make some changes to this figure? See questions in caption of the figure). This may occur because treatments are not applied consistently over the year, either because heat applications are frequently shut off during some seasons such as when snow cover is present or because some heating methods, even if left on throughout the year, are not capable of applying consistent warming year-round (e.g. Clark *et al.*, 2014a,b; Hagedorn *et al.*, 2010). For example, seasonal precipitation patterns can alter the effectiveness of warming treatments, since both infrared heaters and soil cables may fail to achieve the target temperatures during rainstorms(Peterjohn *et al.*, 1993; Hoeppner & Dukes, 2012). Wind has also been shown to alter thermal efficiency of infrared heaters, so if heater capacity is limited, target warming levels may not be reached during windy conditions(Kimball, 2005; Kimball *et al.*, 2008).

Experimental warming effects can also vary across years (Figure 4). This can be due to interactive effects of warming treatments and precipitation, wind, or other aspects of weather that may vary annually, as well as seasonally, as discussed above. Of course, there is bound to be variation in the amount of warming at daily, seasonal, and annual scales as anthropogenic warming progresses, as well. We do not wish to suggest that experimental treatments must be consistent one hundred percent of the time, in order to be relevant. Rather, we wish to call attention to the fact that variations in warming treatments at daily, seasonal, and annual time scales are rarely analyzed explicitly. It is unknown how divergent these annual, seasonal, and daily variations may be from real (i.e. non-experimental) climate patterns. To better understand this potential divergence, we need a detailed comparison of the variation present in climate change experiments to observations in non-experimental settings, as well as to projected future changes given expected greenhouse gas emissions.

Treatments vary in space

In addition to temporal variation, there can be spatial variation in experimental warming effects, such that extrapolation of experimental warming to forecast climate change impacts may not be a straightforward space-for-time substitution (Johnson & Miyanishi, 2008; Jochner *et al.*, 2013). For example, the C3E database contains three 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). We found that the amount of warming may vary by more than one degree among blocks (Figure 4, Table 1S), resulting in lower warming treatments that varied almost 100% of their target temperature and higher warming treatments varying 20%

There are numerous potential causes for these differences in warming levels among blocks, given the same warming treatment. It is possible that fine-scale variation in vegetation, slope, aspect, soil type, or other factors alter wind or soil moisture which in turn affect the thermal efficiency of heaters, or other aspects of the warming treatment, as discussed above(Peterjohn *et al.*, 1993; Kimball, 2005; Kimball *et al.*, 2008; Hoeppner & Dukes, 2012). In addition, fine-scale variation in the amount of sun received by a plot, perhaps due to vegetation cover or slope, may also affect temperature (Rollinson & Kaye, 2015). The observed differences in effective warming among blocks highlights the importance of quantifying temperature, soil moisture, and other climate variables at the plot scale, as well as a need to better understand how fine-scale variation in slope, vegetation, and other factors influence the microclimatic conditions experienced by organisms.

Presumably, there will be spatial variation in future climate change effects, given that warming to date has varied spatially (Stocker *et al.*, 2013)). Accurate extrapolation of climate change experiments may therefore depend on the extent to which experiments encompass a representative amount of existing natural variation (e.g. gradients in slope and aspect) present at the scale at which the extrapolation is being made. More detailed, fine-scaled comparisons of experimental and observational studies are necessary to better understand how this spatial variation in warming affects our interpretation of responses to warming.

Experimental infrastructure alters climate

The experimental structures themselves alter temperature and other important biotic and abiotic variables, in ways that are not generally examined or reported in experimental warming studies. The possible existence of these effects are widely acknowledged, and some studies include 'shams' or 'disturbance controls' in an attempt to account for them. However, the magnitude and implications of structural effects on climate are rarely discussed or interpreted in climate change studies.

To investigate the magnitude of these 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.*, 2014a; ?; Pelini *et al.*, 2011). These were the only studies in our database that included two types of control plots: structural controls (i.e. 'shams' or 'disturbance controls,' which contained all the warming infrastructure, such as soil cables or infrared heating units but with no heat applied) and ambient controls with no infrastructure added (see Supplemental Materials for details). Other studies include only the structural controls (n=3) or only the ambient controls (n=4).

We found that experimental structures altered air and soil temperatures in opposing ways: air temperatures were higher in the structural controls, compared with the ambient air with no structures installed, whereas soil temperatures were lower in the structural controls compared with ambient soil (Figure5). This general pattern was consistent across the different temperature models we fit (mean, minimum, and maximum), although the magnitude varied across seasons (Figure 5), as well as among studies, years, and with ambient temperature (Table XS). Soil moisture was lower in structural controls compared with ambient conditions (Figure 6).

There are several possible mechanisms for the observed differences between ambient and structural controls. Elevated air temperatures in structural controls may be caused by infrastructure surfaces absorbing more solar

radiation. For example, infrastructure materials may have higher heat capacity, higher emissivity, and/or lower albedo than the surrounding ground surface. In addition reduced air flow within the structures may lower evaporative cooling from structural control compared with ambient control plots. Although we could find very little discussion of measured temperature (or other) differences between ambient and structural control plots in most previously published work (e.g. Farnsworth *et al.*, 1995; Pelini *et al.*, 2011; Clark *et al.*, 2014a,b), Clark et al (2014) do mention that "control of the air temperature was less precise, in part due to air scooping on windy days." Marchin et al. (2015) also note that structural controls had mean spring air temperatures about 0.5ÂřC or more above ambient temperatures. As for the differences in soil temperatures, it may be caused by alterations to heat capacity and emissivity in the structural materials compared to surrounding soil. Structures could also interfere with snow accumulation, thereby reducing snowpack and its insulation; this likely plays a bigger role at the Harvard Forest sites (exp04, exp07), where average snowfall is close to four feet, than at Duke Forest, where average snow accumulation is 7.5 inches or less. Peterjohn (1994) also reported cooler temperatures in structural versus ambient control plots, but only at shallow soil depths (4 cm deep, in their study). Similarly, in our analysis we found the greatest difference between soil temperature in sham and ambient controls to be in exp10, one of the two studies (the other was exp07) in which temperature was measured at depths of 2 cm, rather than 15 cm deep (exp03 and exp04).

In addition to the structural effects that we document here on temperature and moisture, experimental structures may alter conditions by reducing light, intercepting precipitation, and altering herbivory and other biotic interactions(Kennedy, 1995; Wolkovich *et al.*, 2012). Most warming experiments to date deal with this by calculating focal response variables relative to ambient control chambers to account for chamber effect(e.g. Marchin *et al.*, 2015). Further documentation and analysis of the effects on abiotic and biotic factors, as well as in depth interpretation of how these effects may alter focal variables, is an important next step for climate change experimentation, particularly if we wish to apply results to forecasting. To date, these side effects are rarely reported or interpreted in climate change experiments, making them "demonic intrusions" in need of âĂIJeternal vigilance" to reduce(Hurlbert, 1984). There will always be some artifacts present in any experiment, no matter how perfectly designed; we argue here for a better understanding of what those artifacts are in climate change experiments, and a fuller interpretation of how, if at all, they may affect translating the results of these experiments into forecasts.

Secondary effects of climate change manipulations

Climate change experiments often seek to manipulate one or two climate variables, such as temperature and precipitation. However, there are likely to be non-target abiotic and biotic factors that are also affected by these manipulations. For example, precipitation treatments typically reduce temperatures in climate change manipulations(Sherry *et al.*, 2007; Rollinson & Kaye, 2012; McDaniel *et al.*, 2014). McDaniel et al. (2014) observed that a twenty percent increase in precipitation reduced mean hourly temperatures by 0.3 degrees Celsius over the course of their two-year experiment. The magnitude of this effect can vary in space and time, however (Figure 3).

In addition, experimental warming typically reduces vapor pressure deficit and soil water content (Figure 7),(e.g. Sherry *et al.*, 2007; Morin *et al.*, 2010; Templer *et al.*, 2016). Of the twelve experiments in the C3E database, ten measured and reported soil moisture, and five measured air and soil temperature in addition to soil moisture. We found that soil moisture was reduced by 0.2 percent, on average, per degree of air warming (Table XS). While active warming experiments rarely manipulate soil moisture directly, soil moisture is unavoidably affected by changing temperatures. Because soil moisture is one of the most fundamental quantities affecting plant physiological functioning (Lambers et al. 2008) and climate system cycles of water and energy (Seneviratne et al. 2010), consideration of its alteration by experimental warming is vital.

(Miriam: I redid your models, to match the structure of the other models in the paper: with site and year

(nested within site) as random effects. This changed the results slightly. Could you work on this paragraph a bit, and add some interpretation with the new model structure and results?)

Warming and precipitation treatments, and their indirect effects on abiotic factors such as soil moisture, can also alter the biotic environment, which in turn can produce additional secondary effects that alter climate. For example, Rollinson et al (2010) reported that tree composition shifted after three years of warming and modified precipitation treatments (Rollinson & Kaye, 2012). These shifts in composition may change competitive dynamics and, in turn, affect resource levels, such as moisture in the soil. In addition, given that warming reduces soil water content, it is likely to affect soil microbial communities, and therefore available nutrients as well. The magnitude of all of these effects are also likely to vary in space and time; some may be transient whereas others may be more permanent.

It can be difficult to tease apart the specific abiotic and biotic drivers climatic conditions in climate change experiments, but understanding the effects of an experimental treatment on these interrelated variables is critical when trying to determine mechanistic explanations for observed responses to warming. Even when experimental artifacts are introduced (such as artificial drying of soils or altered plant composition, due to experimental structures), if these secondary effects are quantified, they can be helpful in understanding how abiotic and biotic factors interact to affect physiological responses. For example, we can learn about the controls on stomatal conductance when the normal covariance between temperature, humidity, and soil moisture is altered.

The widespread presence of unintended secondary effects of climate change manipulations highlight the importance of measuring environmental conditions at the plot level, and of using these measurements in data analysis and interpretation of results. Many climate change experiments (seven of the 12 in the C3E database, for example) model warming and/or precipitation treatments as categorical predictors (and in some cases, orthogonal crossed treatments, when both treatments are included in the experiment, i.e. a traditional repeated measures, three-way ANOVA). The interacting and secondary effects of these manipulations, as well as the plot-level variation in warming effectiveness and effects of experimental structures on temperature and soil moisture that we discuss above, demonstrate a clear need for an alternative modelling approach to fully understand the experimental results. One option is to include the continuous climate data (e.g. mean temperature over the study period) for each plot, as a predictor of the focal response variable, such as phenological state or species density (e.g. Marchin *et al.*, 2015; Pelini *et al.*, 2014). A challenge with this approach is that much of the true variation in the climate is lost through aggregation (e.g. calculating mean annual or seasonal temperature), and chosen method of aggregation affects both the mean and variance of the climate estimate (e.g. Clark *et al.*, 2014b). It may not be obvious which method of aggregation, or which combination of aggregated climate variables, is most appropriate, and this will likely depend on the response variable of interest. In these cases, model selection approaches have been used to identify the climate aggregation method that best explains the focal response(e.g. Morin *et al.*, 2010). Alternatively, a continuous development model can be used to capture the full range of variation present in a climatic variable during the study period (e.g. Clark *et al.*, 2014b).

Biological implications

We have highlighted a suite of factors that complicate interpretations of warming experiments. We argue that these largely unintended alterations are important for scientists to fully understand and report in their research (Figure 8). This is especially important because unintended climate alterations are likely to have biological implications, including for many of the major responses studied in warming experiments. Below, we discuss three example biological responses often studied in climate change experiments (plant phenology, plant growth, and soil respiration), for which indirect effects of treatments may have important implications. We argue that the interpretation of how these responses are affected by climate change is likely to be altered by using fine-scale, measured climate as the explanatory variables (e.g.. plot-level temperature and soil

moisture), rather than by using the intended climate treatments (i.e. categorical comparisons or target warming levels). This is because detailed examination of multiple microclimate variables will allow a more complete understanding of the indirect, as well as direct, effects of applied treatments on a suite of abiotic and biotic drivers of focal responses.

Shifted plant phenology has been a focal response in climate change experiments experiments. Yet understanding exactly what drives shifted phenology may be more complicated than simply comparing shifts to the direct warming effects of the experiment. This is because phenology is likely to be altered in opposing ways by the increased air temperatures—which generally advance phenology (Wolkovich *et al.*, 2012)—and decreased soil moisture—which may delay phenology or at least reduce advancement due to warming (Peñuelas *et al.*, 2004; Craine *et al.*, 2012; Matthews & Mazer, 2016)—characterized by warming treatments. Indeed, these opposing drivers may be responsible for the observed discrepancy between observational and experimental phenology responses to warming (Wolkovich *et al.*, 2012). However, observed effects of precipitation treatments on phenology have been variable to date, and remain poorly understood, perhaps because many previous studies have used the applied treatments, rather than the measured microclimate variables, as predictors in analyses (but see Morin *et al.*, 2010). Climate change experiments that manipulate and measure air or soil temperature and soil moisture levels can be used to estimate effect sizes for these two climate variables and potential interactive effects of them on phenology. In addition, plant phenology responds to minimum temperatures, as well as mean and maxima (Shen *et al.*, 2016; Fu *et al.*, 2016; Piao *et al.*, 2015). This may also play a role in the discrepancy between observational and experimental studies, since diurnal versus night temperatures are affected differently by warming treatments(Shen *et al.*, 2016; Matthews & Mazer, 2016).

Plant growth is also likely to be altered in opposing ways by the increased air temperatures and decreased soil moisture levels in experimentally warmed plots. For example, with warming and decreased vapor pressure deficit, stomata closure may reduce sapflow and growth (Templer *et al.*, 2016). Even small shifts in temperature may have a big effect, since the photosynthetic response to temperature is nonlinear (Berry & Bjorkman, 1980). Climate change experiments offer the opportunity to get these (and other) physiological measurements from a wide range of temperature conditions, which is essential for improving accuracy of ecosystem models and their forecasts.

Direct and indirect effects of climate change experiments are also likely to affect soil respiration in ways that may alter carbon dynamics and net mineralization and therefore have other cascading effects that may or may not be realistic (Del Toro *et al.*, 2015). . For example, a meta-analysis based on studies using different warming techniques found that soil warming significantly increased soil respiration rate, net nitrogen mineralization rate and plant productivity, especially in temperate forest ecosystems (Rustad *et al.*, 2001). The amount of increase in these important ecosystem properties is likely to vary dramatically, depending on experimental methods. Warming treatments that are not applied year around (e.g. Clark *et al.*, 2014a,b) may result in effects that differ from those under consistent warming, since microbial communities that have strong effects on nutrient availability and soil moisture may be differentially affected by warming during summer versus during winter, when temperatures are close to critical cold thresholds and when the frequency of freeze-thaw events are likely to be altered(Rivkina *et al.*, 2000; McDaniel *et al.*, 2014). Furthermore, nutrient availability will then alter plant growth, a common focal response in climate change experiments: even a relatively small amount of newly available nitrogen from the soil could result in a substantial increase in carbon storage in woody tissues, especially if it becomes available during critical growth periods for plants. (Fatichi *et al.*, 2014) YANN- Is this the correct citation? In addition, warming of soil versus air is likely to result in different effects on plant-soil microbe interactions, since soil warming experiments may increase net mineralization in early spring or winter, but plant phenology would be unlikely to advance in concert, unless additional air warming were applied (Du & Fang, 2014). These kinds of trophic mismatches may be present in climate change experiments, even if they are not expected to occur under natural warming (Kharouba *et al.*, 2015).

Other biotic interactions are also likely to be affected by direct and indirect affects of climate change treatments. For example, Hoeppner and Dukes (2012) found that rodent disturbance varied by warming treatment (as well as year) in their climate change experiment. Insect diversity and community structure can change

with active warming, and may have secondary effects on plants or other interacting organisms(Pelini *et al.*, 2014; Diamond *et al.*, 2016). A critical question is the extent to which these shifts in biotic interactions (and their effects on focal responses) are accurate forecasts of future shifts that are likely to occur with climate change, or due to side-effects that are unlikely to occur outside of experimental systems (Diamond *et al.*, 2013).

Recommendations for future climate change experiments

Climate change experiments provide invaluable information about biological responses to climate change, yet our results highlight that we do not fully explore the ways in which these climate change experiments are actually altering climate. These complications should not suggest that experimental climate change studies offer little value. Instead, we believe these complications and the related climate data provide the foundation to designing better experiments and gaining the most knowledge and utility from existing experiments. Below we describe recommendations to improve implementation, interpretation, and communication of climate change experiments in the future.

Include both structural and ambient controls and collect, use, and report data collected within them. Fewer than half of the studies in our C3E database included these two control types and monitored climate within them (5 out of 12); all experiments showed significant effects of shams. As far as we know, only one of the studies in the database monitored both climate and focal biological response variables in both types of controls, however (e.g. plant phenology **?**). Future monitoring of all climate and biological variables in both control types will enable scientists to tease apart mechanisms due to experimental design from mechanisms due to actual shifts in climate.

Maximize the length of climate change experiments by running them for as long as possible. This will allow study of how inter-annual variations interact with climate change treatments, especially when looking at non-linear and mulch-year processes such as phenology. It will also allow us to understand how long-term responses may differ from transient ones (Franklin, 1989; Giasson *et al.*, 2013).

Collect fine-scale climate data, at least twice daily, and ideally hourly, to allow for minimum and maximum values, as well as critical thresholds, such as the number and duration of freeze-thaw events and accumulated chilling hours, to be to be analyzed and interpreted, in addition to mean values (McDaniel *et al.*, 2014). In assessing the frequency of measurements, it is critical to consider the chronobiology of the event and organisms of interest. For ants, this might mean that temperatures be monitored at the frequency of minute (Aaron: do you have a citation for this); for bacteria, even more frequently.

Analyze continuous, plot-level climate variables, rather than discrete treatments (i.e. ANOVAs). There can be substantial variation among plots in effects of warming and precipitation treatments (Figure 3)). Furthermore, these two treatments are not independent: precipitation treatments alter the effectiveness and warming, and warming treatments alter available moisture levels. Analyzing climate in this way will allow much more in-depth understanding of the drivers and effects of variation in temperature and moisture.

Consult observational records and forecasts to design relevant manipulations. If the goal is to mimic future climate conditions, careful consultation of climate change projections for the study region, as well as historical data, can aid in selection of warming and precipitation treatment methods that most closely mimic anticipated changes. When it is not possible to match anticipated changes in climate, studies should report how imposed treatments compared to projected changes. In addition, the timing of the imposed treatments should be carefully considered, and ideally should match forecasts. If it is not possible to apply continuous treatments throughout the study, the seasonality and timing of treatments should be explicitly reported and the climate should be monitored throughout, even if no manipulations are implemented. In addition to designing experiments that shift mean climatic conditions, scientists should be thoughtful about what other aspects of microclimate are being altered, such as the minimum, maximum, variance, distribution, and

critical threshold values. In particular, extremes can often have out-sized effects compared to shifts only in the mean state (Ben could you add a citation here).

Publish high quality, usable data and metadata such that data can easily be shared and used by others. In the metadata, report the number and cause of missing data points for climate, especially those collected in warming treatments. (For example, are data missing because the heaters failed, or because sensors failed) Report the timing of applied warming treatments (i.e. exact start and end dates, within and across years), as well as variations in daytime and nighttime and seasonal variations in climate variables. Given that experimental in situ active 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 more in-depth understanding of these valuable data.

Consider implementing and following community standards for reporting climate data When studying biological implications of a global challenge as large as climate change, it will facilitate progress if we can design, run, and report experiments in such a way that we can eventually create a global data set. This recommendation stems from our work gathering and analyzing data from many climate change experiments. We found that studies report a diverse range of climate variables, collected in different ways (e.g. in the the C3E database, soil temperature was collected at depths ranging from 2 to 25 cm and soil moisture was collected at depths ranging from 8 to 30 cm, using different units and methods). It has been difficult to synthesize these data in a comprehensive way that can fully address important questions, and it will be a challenge to tease apart whether variable findings are due to methodological differences, to measurement error, or to true variation in biological responses. Question for everyone (Lizzie/Ben/Miriam/Ann Marie/Aaron/Yann/Isabelle/Jeff/Christy): I'm not sure if i want to keep this recommendation- perhaps its the part of me with libertarian/anti-top-down tendencies (although i can certainly see the value in this). What do others think?

Documenting biological impacts of climate change has over a 20 year history in ecology today. During this time, in situ field experiments have been critical in making the mechanistic link between warming and a number of major biological impacts, such as changes in productivity, soil respiration, the phenology of plant and animals, and shifted community and ecosystem dynamics. Yet, as climate change across the globe continues with projected warming likely to exceed 2 degrees Celsius over the next 80 years (Stocker *et al.*, 2013), ecologists are challenged to not only document impacts but make quantitative robust predictions. Our ability to meet this challenge requires building on the data from current and past experiments to best understand how changes in climate alter ecological processes, and to build better experiments in the future. As a first step, we have compiled the first database of fine-scale climate data from multiple warming experiments and shown how time, space and artifacts may hinder simple interpretations of climate change experiments. The next steps require the ecological community to build on these data and their findings to develop and use new approaches in future experiments. This will provide more accurate estimates of altered climate in these experiments and in turn, more accurate estimates of critical biological changes.

References

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

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

Thomas, C. D. *et al.* Extinction risk from climate change. *Nature* **427**, 145–148 (2004). PT: J.

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

Field, C. B., Lobell, D. B., Peters, H. A. & Chiariello, N. R. Feedbacks of terrestrial ecosystems to climate change*. *Annu. Rev. Environ. Resour.* **32**, 1–29 (2007).

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

Urban, M. C., Tewksbury, J. J. & Sheldon, K. S. 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 (2012). PT: J; NR: 38; TC: 1; J9: P ROY SOC B-BIOL SCI; PG: 9; GA: 925RT; UT: WOS:000302779600025.

Stocker, T. *et al.* IPCC, 2013: climate change 2013: the physical science basis. Contribution of working group I to the fifth assessment report of the intergovernmental panel on climate change (2013).

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

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

Williams, J. W., Jackson, S. T. & Kutzbacht, J. E. 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 (2007). PT: J.

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

Ibanez, I. *et al.* Predicting biodiversity change: Outside the climate envelope, beyond the species-area curve. *Ecology* **87**, 1896–1906 (2006). PT: J.

Swab, R. M., Regan, H. M., Keith, D. A., Regan, T. J. & Ooi, M. K. J. Niche models tell half the story: spatial context and life-history traits influence species responses to global change. *Journal of Biogeography* **39**, 1266–1277 (2012). PT: J; NR: 63; TC: 0; J9: J BIOGEOGR; PG: 12; GA: 961GP; UT: WOS:000305452500006.

Box, G. E., Hunter, W. G., Hunter, J. S. *et al.* Statistics for experimenters (1978).

Gelman, A. *Experimental reasoning in social science experiments*, chap. 7, 185–195 (New Haven, CT: Yale University Press, 2014).

Shaver, G. R. *et al.* 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 (2000).

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

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

Cleland, E. E., Chiariello, N. R., Loarie, S. R., Mooney, H. A. & Field, C. B. 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 (2006). LR: 20140908; JID: 7505876; 0 (Soil); 142M471B3J (Carbon Dioxide); N762921K75 (Nitrogen); OID: NLM: PMC1560087; 2006/09/05 [aheadofprint]; ppublish.

Sherry, R. A. *et al.* Divergence of reproductive phenology under climate warming. *Proceedings of the National Academy of Sciences of the United States of America* **104**, 198–202 (2007). LR: 20140907; JID: 7505876; OID: NLM: PMC1713188; 2006/12/20 [aheadofprint]; ppublish.

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

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

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

Reich, P. B. *et al.* Geographic range predicts photosynthetic and growth response to warming in co-occurring tree species. *Nature Clim.Change* **5**, 148–152 (2015). URL `http://dx.doi.org/10.1038/nclimate2497`, 2015/02//print.

Gruner, D. S. *et al.* Effects of experimental warming on biodiversity depend on ecosystem type and local species composition. *Oikos* (2016).

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

Kimball, B. A. *et al.* Infrared heater arrays for warming ecosystem field plots. *Global Change Biology* **14**, 309–320 (2008).

Wolkovich, E. M. *et al.* Warming experiments underpredict plant phenological responses to climate change. *Nature* **485**, 494–497 (2012). PT: J; UT: WOS:000304344500041.

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

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

McDaniel, M. *et al.* Microclimate and ecological threshold responses in a warming and wetting experiment following whole tree harvest. *Theoretical and applied climatology* **116**, 287–299 (2014).

Shen, M. *et al.* Strong impacts of daily minimum temperature on the green-up date and summer greenness of the Tibetan Plateau. *Global change biology* (2016).

Matthews, E. R. & Mazer, S. J. 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 (2016).

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

Hagedorn, F. *et al.* Short-term responses of ecosystem carbon fluxes to experimental soil warming at the Swiss alpine treeline. *Biogeochemistry* **97**, 7–19 (2010).

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

Johnson, E. A. & Miyanishi, K. Testing the assumptions of chronosequences in succession. *Ecology Letters* **11**, 419–431 (2008).

Jochner, S., Caffarra, A. & Menzel, A. Can spatial data substitute temporal data in phenological modelling? A survey using birch flowering. *Tree physiology* **33**, 1256–1268 (2013).

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

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

Pelini, S. L. *et al.* Heating up the forest: openâĂŘtop chamber warming manipulation of arthropod communities at Harvard and Duke Forests. *Methods in Ecology and Evolution* **2**, 534–540 (2011).

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

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

Hurlbert, S. H. Pseudoreplication and the design of ecological field experiments. *Ecological monographs* **54**, 187–211 (1984).

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

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

Pelini, S. *et al.* Geographic differences in effects of experimental warming on ant species diversity and community composition. *Ecosphere* **5**, 1–12 (2014).

Peñuelas, J. *et al.* Complex spatiotemporal phenological shifts as a response to rainfall changes. *New Phytologist* **161**, 837–846 (2004).

Craine, J. M., Wolkovich, E. M., Towne, E. G. & Kembel, S. W. Flowering phenology as a functional trait in a tallgrass prairie. *New Phytologist* **193**, 673–682 (2012). PT: J; TC: 3; UT: WOS:000298984900013.

Fu, Y. H. *et al.* Three times greater weight of daytime than of night-time temperature on leaf unfolding phenology in temperate trees. *New Phytologist* (2016).

Piao, S. *et al.* Leaf onset in the northern hemisphere triggered by daytime temperature. *Nature communications* **6** (2015).

Berry, J. & Bjorkman, O. Photosynthetic response and adaptation to temperature in higher plants. *Annual Review of Plant Physiology* **31**, 491–543 (1980).

Del Toro, I., Ribbons, R. R. & Ellison, A. M. Ant-mediated ecosystem functions on a warmer planet: effects on soil movement, decomposition and nutrient cycling. *Journal of Animal Ecology* **84**, 1233–1241 (2015).

Rustad, L. *et al.* A meta-analysis of the response of soil respiration, net nitrogen mineralization, and aboveground plant growth to experimental ecosystem warming. *Oecologia* **126**, 543–562 (2001).

Rivkina, E., Friedmann, E., McKay, C. & Gilichinsky, D. Metabolic activity of permafrost bacteria below the freezing point. *Applied and Environmental Microbiology* **66**, 3230–3233 (2000).

Fatichi, S., Leuzinger, S. & Koerner, C. Moving beyond photosynthesis: from carbon source to sink-driven vegetation modeling. *New Phytologist* **201**, 1086–1095 (2014).

Du, E. & Fang, J. Linking belowground and aboveground phenology in two boreal forests in Northeast China. *Oecologia* **176**, 883–892 (2014).

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

Diamond, S. E. *et al.* Climate warming destabilizes forest ant communities. *Science Advances* **In press** (2016).

Diamond, S. E. *et al.* Using physiology to predict the responses of ants to climatic warming. *Integrative and comparative biology* **53**, 965–974 (2013).

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

Giasson, M.-A. *et al.* Soil respiration in a northeastern US temperate forest: a 22-year synthesis. *Ecosphere* **4**, 1–28 (2013).

Figures
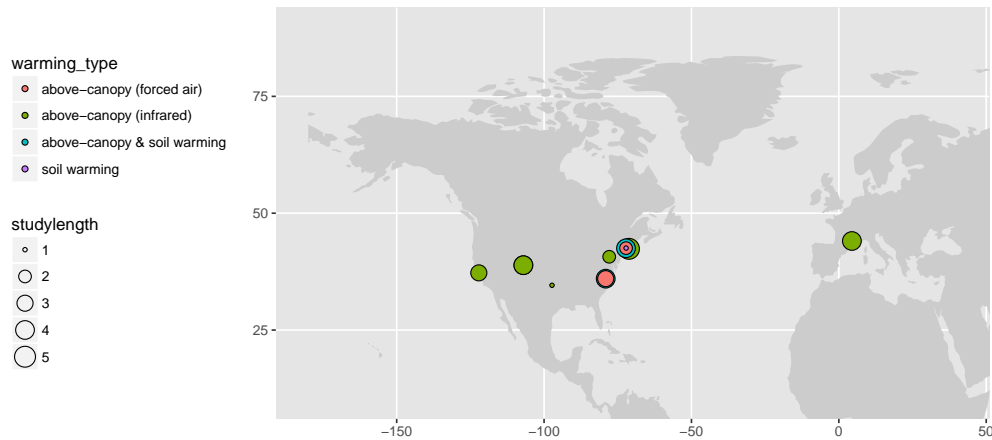
Figure 1: DRAFT Map of experimental sites in C3E datbase. Question for everyone (Lizzie/Ben/Miriam/Ann Marie/Aaron/Yann/Isabelle/Jeff/Christy): Does anyone want to volunteer to make a beautiful new version?
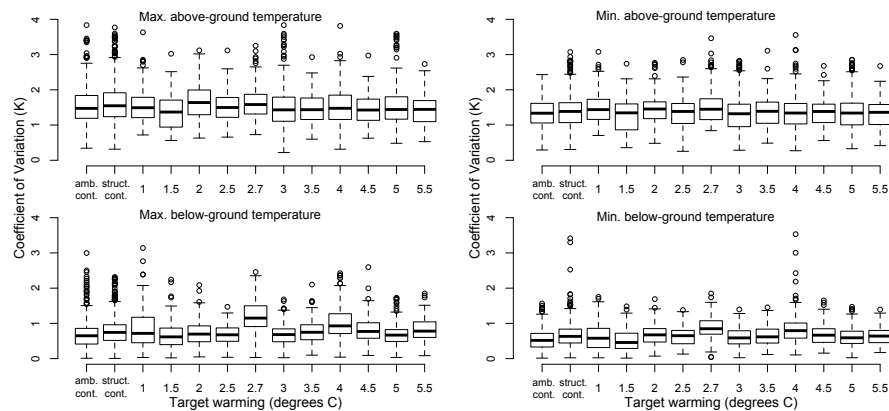
Figure 2: Ambient controls have reduced variation, compared with structural controls and nearly all warming treatments. I'm not happy with this figure- I've tried so many different ways of showing the (small but) significant differences in temperature range/variance among ambient controls, structural constrols and warming treatments, in addition to the statistical analyses described in the text and supplement. Question for everyone (Lizzie/Ben/Miriam/Ann Marie/Aaron/Yann/Isabelle/Jeff/Christy): I need help with ideas for other ways to show this, or thoughts on if the point should be dropped, since the differences are minor and hard to see.
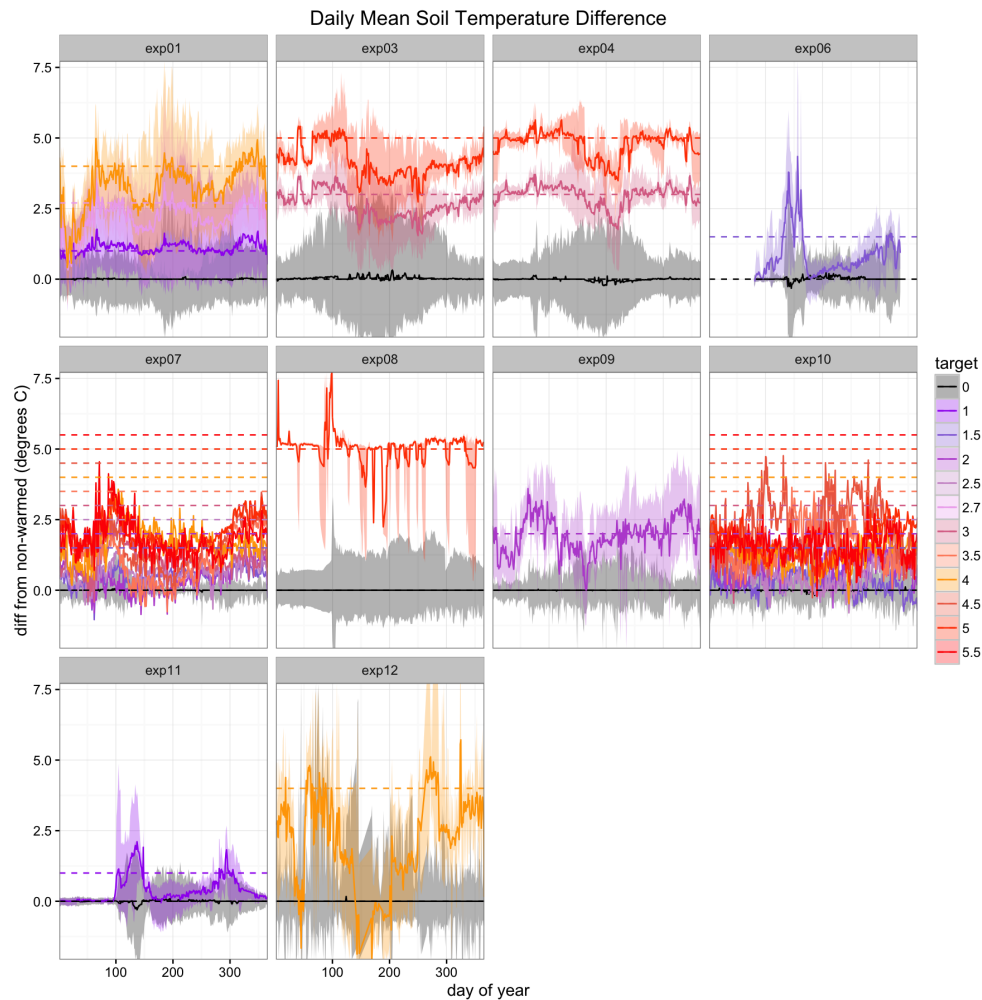
Figure 3: Time series of deviations from mean soil temperature over one year, in control (black line) and warming treatments with various target warming levels at 10 study sites. Christy: could you add the mean annual temperature somehow to each panel- just so that we can see the variation among sites? Also, I am wondering if you could modify the code to make these estimates for each year of a study, as well as overall. I also was not quite sure if this represents a summary of deviations across all years, if studies were multiple years, or just one year? Please write a brief summary of these analyses-probably for the supplemental.
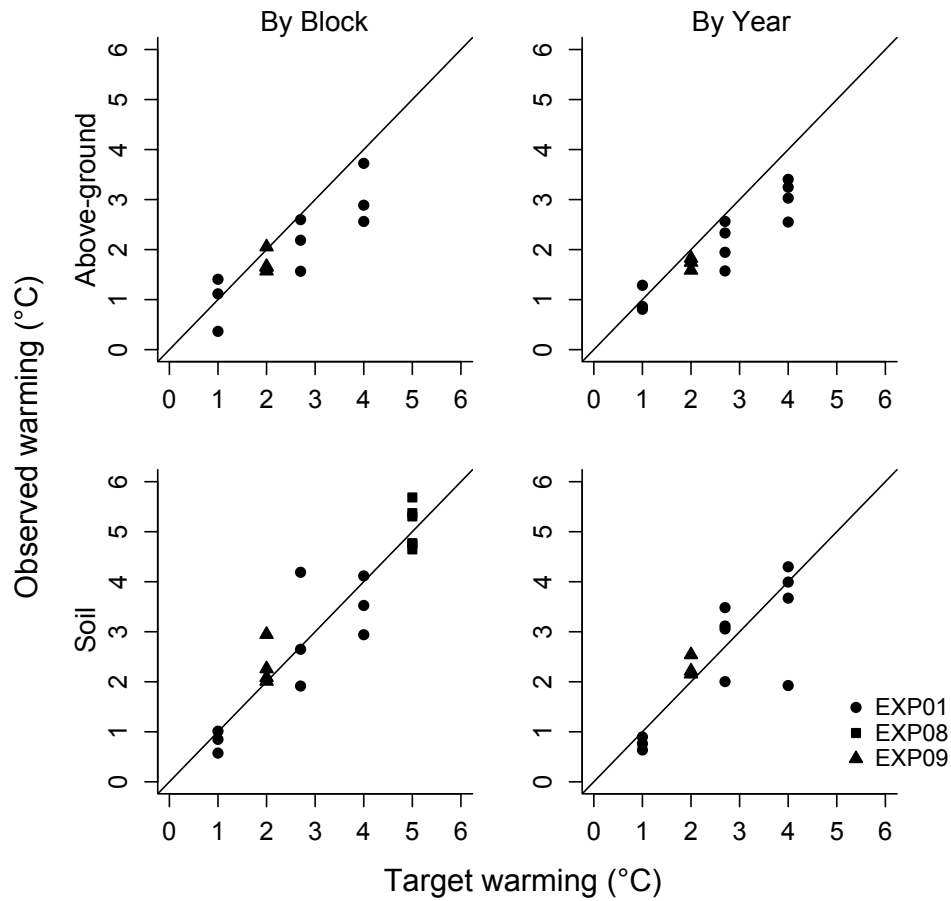
Figure 4: The amount of observed warming (i.e. the difference between treatment and control plots, within each block) varies among blocks (left panels), as well as among years (right panels). Black line represents a 1 to 1 ratio (i.e. when observed warming is exactly equal to target warming). See Tables XS and XS for statistical differences.
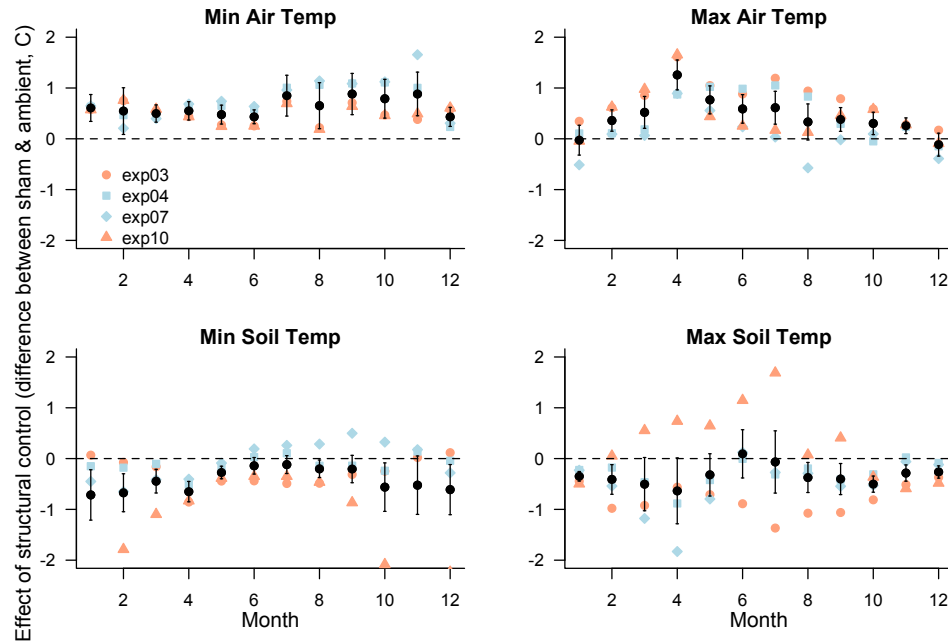
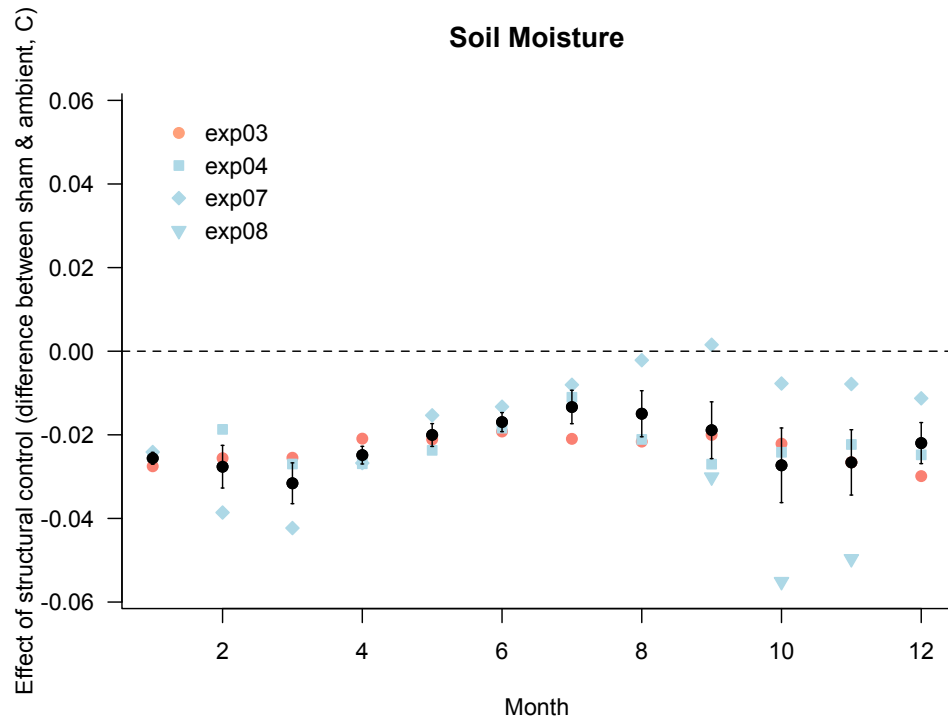Figure 5: Air temperatures were higher, whereas soil temperatures were lower, in general, in structural controls compared with ambient controls (i.e. with no control chambers or warming infrastructure in place). We show fixed effects (in black) from a mixed effects model that accounts for differences in experimental design and other factors among sites with mixed effects model including site and year (nested within site) as random effects (see Table XS, Supplemental Materials for details). Site-level random effects are shown by symbols in blue (for those at Harvard) and pink (those at Duke).

Figure 6: Soil moisture was lower in structural controls compared with ambient controls (i.e. with no control chambers or warming infrastructure in place). We show fixed effects (in black) from a mixed effects model that accounts for differences in experimental design and other factors among sites with mixed effects model including site and year (nested within site) as random effects (see Table XS, Supplemental Materials for details). Site-level random effects are shown by symbols in blue (for those at Harvard) and pink (those at Duke).
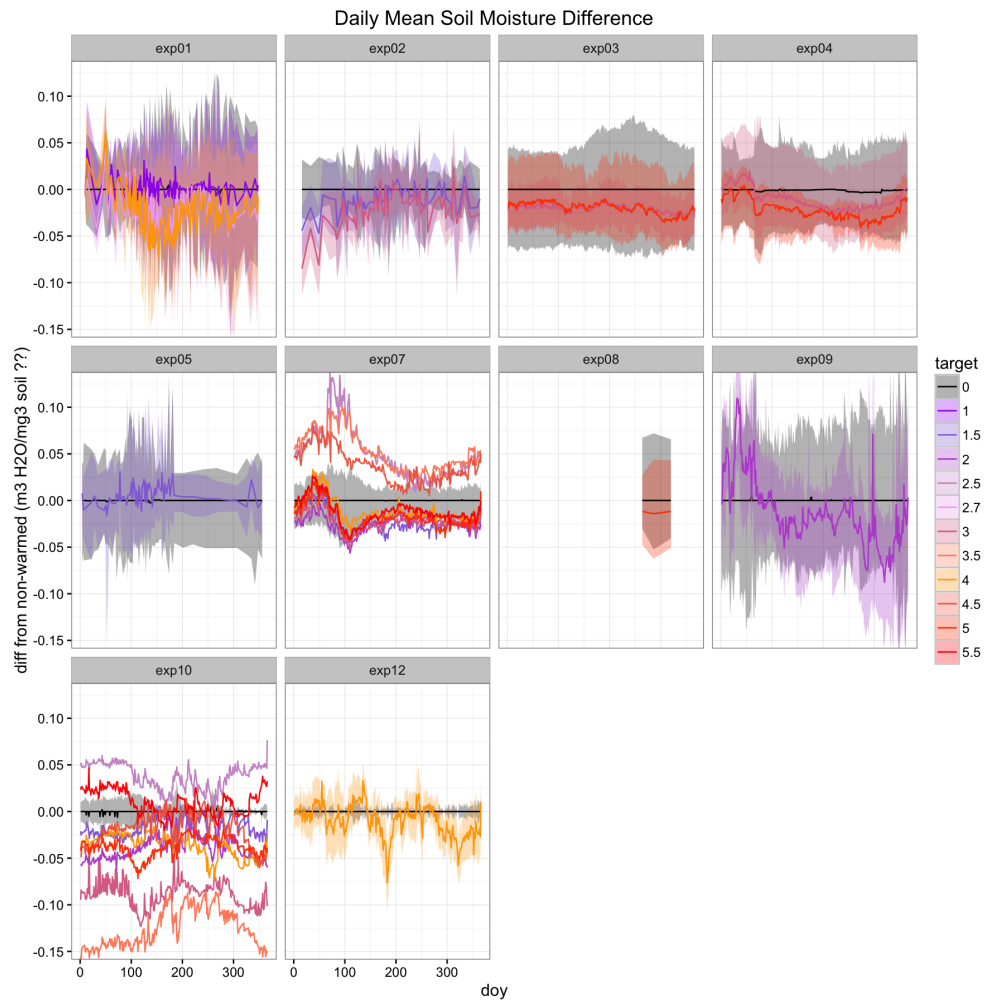
Figure 7: Time series of deviations from mean daily soil moisture in control (black line) and warming treatments with various target warming levels at 10 study sites.
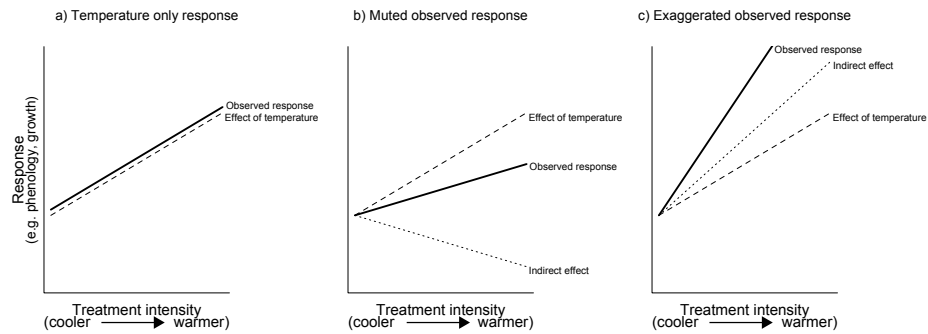
Figure 8: Compared to direct responses to temperature alone (a), experimental warming may cause biological responses to be muted (b)or exaggerated (c), when indirect effects of experimental warming are also drivers of focal responses. For example, phenology, such as budburst which usually advances with warming, may appear to be less sensitive to warming in experiments versus observational studies (Wolkovich *et al.*, 2012) because experimental warming reduces soil moisture and drying often delays phenology.
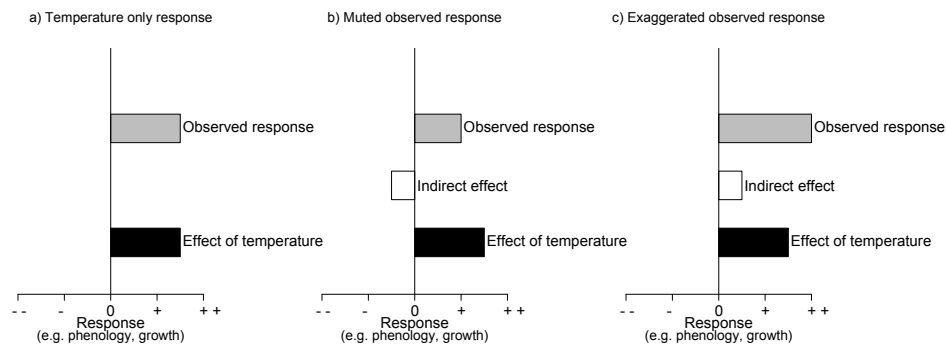

Figure 9: Question for everyone (Lizzie/Ben/Miriam/Ann Marie/Aaron/Yann/Isabelle/Jeff/Christy): This is an alternative to the above figure- which do you prefer? Compared to direct responses to temperature alone (a), experimental warming may cause biological responses to be muted (b)or exaggerated (c), when indirect effects of experimental warming are also drivers of focal responses. For example, phenology, such as budburst which usually advances with warming, may appear to be less sensitive to warming in experiments versus observational studies (Wolkovich *et al.*, 2012) because experimental warming reduces soil moisture and drying often delays phenology.