

How do climate change experiments alter plot-scale climate?

A.K. 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^{1,8,13,j}

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

²Tufts University, Medford, Massachusetts 02155, USA

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

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

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

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

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

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

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

¹⁰Center for Tree Science, The Morton Arboretum, Lisle, Illinois 60532, USA

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

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

¹³Forest & Conservation Sciences, Faculty of Forestry, University of British Columbia,
Vancouver, BC, Canada

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

^bisabelle.chuine@cefe.cnrs.fr

^cbc9z@ldeo.columbia.edu

^djsdukes@purdue.edu

^eaellison@fas.harvard.edu

^fmjohnston@g.harvard.edu

^ganne.panetta@colorado.edu

^hcrollinson@mortonarb.org

ⁱyann.vitasse@wsl.ch

^je.wolkovich@ubc.ca

October 5, 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. AKE and EMW conceived of the idea for the literature review, database compilation, and related Radcliffe Exploratory Seminar. AKE compiled the datasets; AKE and CRR analyzed the data and created the figures; AKE 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 4,933

Number of references 97

Number of figures 5

Number of text boxes 2

Number of words in Box 1

Number of words in Box 2 570

¹ Abstract

² To understand and forecast biological responses to climate change, scientists frequently use field experiments
³ that alter temperature and precipitation. Climate manipulations can manifest in complex ways, however,
⁴ challenging interpretations of biological responses. We reviewed publications from active-warming experi-
⁵ ments to compile a database of daily plot-scale climate data from 15 experiments that use forced air, infrared
⁶ heaters, or soil cables to warm plots. We find that the common practices of analyzing primarily mean
⁷ changes among treatments and analyzing treatments as categorical variables (e.g., warmed verses unwarmed)
⁸ masks important variation in treatment effects over space and time. Our synthesis showed that measured
⁹ mean warming, in plots with the same target warming, differed by 1.6 °C, on average, with high variation—
¹⁰ maximum differences ranged from 0.03 - 4.87 °C across studies, which include variation due to warming
¹¹ design and myriad ecosystem attributes. Warming treatments produce non-temperature effects as well, such
¹² as soil drying. The combination of these direct and indirect effects is complex, with important biological
¹³ consequences. We show one such consequence with a case study of plant phenology, in which accounting
¹⁴ for drier soils with warming triples the estimated sensitivity of budburst to temperature. Our synthesis
¹⁵ yields recommendations for future analyses, experimental design, and data sharing to improve our mecha-
¹⁶ nistic understanding of what drives variation across studies, and thus improve the utility of climate change
¹⁷ experiments to accurately identify and forecast species' responses.

¹⁸ Introduction

¹⁹ Climate change is dramatically altering earth's biota, shifting the physiology, distribution, and abundance
²⁰ of organisms, with cascading community, ecosystem, and climate effects (Shukla & Mintz, 1982; Cox *et al.*,
²¹ 2000; Thomas *et al.*, 2004; Parmesan, 2006; Field *et al.*, 2007; Sheldon *et al.*, 2011; Urban *et al.*, 2012). Much
²² uncertainty exists about how particular individuals, populations, species, communities, and ecosystems will
²³ respond as warming becomes more extreme (Thuiller, 2004; Friedlingstein *et al.*, 2014). Predicting biolog-
²⁴ ical responses to current and future climate change—and their feedbacks to earth's climate and ecosystem
²⁵ services—is one of the most significant challenges facing ecologists today.

²⁶ Two common approaches for understanding biological effects of climate change are observational studies,
²⁷ which correlate recorded biological patterns with measured climate, and process-based modeling; yet these

28 approaches are insufficient for several reasons. Observational studies and correlative models cannot disen-
29 tangle the causal effects of warming (one aspect of climate) from other factors that have also changed over
30 time, such as successional stage or land use. In addition, models based on correlative data may fail to make
31 useful predictions for future conditions that fall outside the range of historical variability (e.g., Pearson &
32 Dawson, 2004; Hampe, 2004; Ibanez *et al.*, 2006; Swab *et al.*, 2012; Chuine *et al.*, 2016). Climate change will
33 yield warmer temperatures than the previous 150 years, and possibly warmer than at any time in the last
34 2000 years (Ohlemüller *et al.*, 2006; Williams & Jackson, 2007; Williams *et al.*, 2007; Stocker *et al.*, 2013).
35 Process-based models overcome some of these challenges through inclusion of explicit mechanistic relation-
36 ships between climate and biological outcomes. However, they are limited by the processes they include (i.e.,
37 our understanding of mechanism), as well as by the data available to parameterize those processes (Moorcroft,
38 2006; Kearney & Porter, 2009).

39 Experimental data from field-based climate change experiments are crucial to fill these knowledge gaps and
40 determine mechanistic links between climate change and biological responses. Experiments can quantify
41 biological responses to different levels of climate change, and can create the “no-analog” climate scenarios
42 forecasted for the future, particularly when they employ active-warming methods, such as forced air heaters,
43 soil warming cables, or infrared heaters (Shaver *et al.*, 2000; Williams *et al.*, 2007; Aronson & McNulty,
44 2009). In addition, active-warming can be combined with precipitation manipulations (e.g., snow removal,
45 water additions or reductions), offering the ability to assess individual and interactive effects of temperature
46 and precipitation, separate from other environmental changes (e.g., Price & Waser, 1998; Cleland *et al.*,
47 2006; Sherry *et al.*, 2007; Rollinson & Kaye, 2012). Compared with indoor growth-chamber experiments,
48 field-based experiments offer the possibility of preserving important but unknown or unquantified feedbacks
49 among biotic and abiotic components of the studied systems.

50 With climate change experiments, ecologists often aim to test hypotheses about how projected warming
51 will affect species’ growth, survival, and future distributions (Dukes & Mooney, 1999; Hobbie *et al.*, 1999;
52 Morin *et al.*, 2010; Pelini *et al.*, 2011; Chuine *et al.*, 2012; Reich *et al.*, 2015; Gruner *et al.*, 2017). Recent
53 research suggests, however, that climate manipulations may not always alter plot-scale climate (hereafter,
54 microclimate) in ways that are consistent with observed changes over time (Wolkovich *et al.*, 2012; Menke
55 *et al.*, 2014; Polgar *et al.*, 2014; Andresen *et al.*, 2016). For extrapolation of experimental findings to the real
56 world, we need detailed assessments of how active-warming experiments alter the microclimate conditions
57 experienced by organisms, and the extent to which these conditions are similar to current field conditions or

58 anticipated climate change.

59 Here, we investigate the complex ways that active-warming treatments alter microclimate, both directly and
60 indirectly, across multiple studies. The qualitative challenges and opportunities of climate change experiments
61 have been summarized previously (e.g., De Boeck *et al.*, 2015) and effects of these manipulations on some
62 aspects of microclimate have been published for individual sites (e.g., Harte *et al.*, 1995; McDaniel *et al.*,
63 2014b; Pelini *et al.*, 2011). However, our quantitative meta-analysis allows us to examine trends across sites
64 and warming designs (Box 1), and make recommendations based on this information. Using plot-level daily
65 microclimate data from 15 active-warming experiments (yielding 59 experiment years and 14,913 experiment
66 days; Table S1), we show the direct and indirect ways that experimental manipulations alter microclimate.
67 We use a case study of spring plant phenology to demonstrate how analyses that assume a constant warming
68 effect and do not include non-temperature effects of warming treatments on biological responses lead to
69 inaccurate quantification of plant sensitivity to temperature shifts. Finally, we synthesize our findings to
70 make recommendations for future analysis and design of climate change experiments (Box 2).

71 MicroClimate from Climate Change Experiments (MC3E) database

72 To investigate how climate change experiments alter microclimate, we first identified published, active-
73 warming field experiments, many of which included precipitation manipulations. We focused on *in situ*
74 active-warming manipulations because recent analyses indicate that active-warming methods are the most
75 controlled and consistent methods available for experimental warming (Kimball, 2005; Kimball *et al.*, 2008;
76 Aronson & McNulty, 2009; Wolkovich *et al.*, 2012). We do not include passive-warming experiments because
77 they have been analyzed extensively already and are known to have distinct issues, including reduction in
78 wind, overheating, and great variation in the amount of warming depending on irradiance and snow depth
79 (Marion *et al.*, 1997; Shaver *et al.*, 2000; Wolkovich *et al.*, 2012; Bokhorst *et al.*, 2013, see also Table S2).

80 We carried out a full literature review to identify potential active-warming field experiments to include in
81 the database. We followed the methods and search terms of Wolkovich *et al.* (2012) for their Synthesis of
82 Timings Observed in iNcrease Experiments (STONE) database (Wolkovich *et al.*, 2012), but restricted our
83 focus to active-warming experiments. Further, because our goal was to tease out variation in microclimate
84 (including temperature and soil moisture), we focused on warming studies that included both/either multiple

85 levels of warming and/or precipitation treatments. These additional restrictions constrained the list to 11
86 new studies published after the STONE database, as well as six of the 37 studies in the STONE database.
87 We contacted authors to obtain daily microclimate and phenological data for these 17 studies and received
88 data (or obtained publicly available data) for 10 of them, as well as datasets from five additional sites offered
89 or suggested to us over the course of our literature review and data analysis. The daily temperature and
90 soil moisture data from these 15 experiments comprise the MicroClimate from Climate Change Experiments
91 (MC3E) database (Figures 1 and S1, Table S1), which is available at KNB (Ettinger & Wolkovich, 2018). We
92 examined how these experiments altered microclimate, using mixed-effects models that allowed for inherent
93 differences among studies (through a random effect of study on the intercept), while also estimating various
94 across-study effects, such as degree of warming or warming type.

95 **Complexities in interpreting experimental climate change**

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

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

113 ecosystems. When possible, we compare and contrast these factors across different study methodologies, such
114 as infrared warming versus forced air chambers and constant wattage versus feedback control, because effects
115 on microclimate may vary across these different methodologies (Box 1).

116 **Effects on microclimate vary over time and space**

117 Reporting only the mean temperature difference across the duration of a warming study masks potentially
118 important temporal variation in temperature among treatments (compare Figure 2 to Figure S2). Using
119 the MC3E database, we found that active-warming reduces the range of above-ground daily temperature
120 by 0.37°C per $^{\circ}\text{C}$ of target warming (Table S3, see also Table S1, which details the different methods used
121 to measure and warm temperatures). Active-warming decreased above-ground daily temperature range by
122 differentially affecting maximum and minimum temperatures: warming increased daily minima by 0.81°C per
123 $^{\circ}\text{C}$ of target warming, but only increased daily maxima by 0.48°C per $^{\circ}\text{C}$ of target warming (Table S3). These
124 effects varied by site (Table S3), but we found no clear patterns by warming type (e.g., infrared versus forced
125 air). Soil daily temperature range was not affected by experimental warming, as warming altered minimum
126 and maximum daily temperatures similarly (Table S4).

127 We observed strong seasonal and annual variations in the effects of experimental warming (Figures 1, 2, Table
128 S5). Warming generally appears close to targets in winter and early spring, and farthest below targets in
129 summer (day of year 150-200, when evapotranspiration within a robust plant canopy can dissipate energy
130 and act to cool vegetation surfaces), though patterns differ among sites (Figure 1). The variation in warming
131 effectiveness may be driven by interactions between warming treatments and daily, seasonal, and annual
132 weather patterns, since the magnitude of warming can vary as weather conditions change. Both infrared
133 heaters and soil cables fail to achieve target temperature increases during rainstorms (Peterjohn *et al.*, 1993;
134 Hoeppner & Dukes, 2012) and with windy conditions (Kimball, 2005; Kimball *et al.*, 2008). Differences
135 between target and actual warming are likely to be particularly great for studies employing constant wattage,
136 rather than feedback control (Box 1, Figure 2). In addition, treatments are often applied inconsistently within
137 or across years. Heat applications are frequently shut off during winter months, and some heating methods,
138 even if left on throughout the year, do not warm consistently (e.g., Clark *et al.*, 2014a,b; Hagedorn *et al.*,
139 2010).

140 Treatment effects also vary spatially, further complicating interpretation of climate change experiments. The

141 MC3E database contains six studies that used blocked designs, allowing us to examine spatial variation in the
142 amount of warming (i.e., the difference between treatment and control plots within a block). These studies
143 include two infrared with feedback control, three infrared with constant energy input, and one soil warming
144 cable with feedback control experiments. We found that the amount of observed warming frequently varied
145 by more than 1°C (mean= 1.6°C, maximum = 4.9°C) among blocks (Figure 2, Table S6). This variation in
146 warming is substantial, as it is equivalent to the target warming treatment for many studies, and appears
147 to vary substantially among sites, which differ in warming methodologies and environmental characteristics,
148 though low sample sizes make disentangling the effect of warming method difficult (Box 1). The differences in
149 warming among blocks may be caused by fine-scale variation in vegetation, slope, aspect, soil type, or other
150 factors that can alter wind or soil moisture, which in turn affect warming (Peterjohn *et al.*, 1993; Kimball,
151 2005; Kimball *et al.*, 2008; Hoeppner & Dukes, 2012; Rollinson & Kaye, 2015).

152 Of course, identical experimental treatments across space and time are neither necessary, nor realistic, for
153 robust analysis of experimental results and forecasting. Indeed, the spatial and temporal variation we re-
154 port could improve and refine models, and—at least in some regions—may be consistent with contemporary
155 patterns of climate change (Stocker *et al.*, 2013). Taking advantage of this variation, though, requires under-
156 standing and reporting it (e.g., Milcu *et al.*, 2016). However, because fine-scale and temporal variations in
157 warming treatments are rarely analyzed explicitly with ecological data, the implications for interpretation of
158 experimental findings are unclear.

159 Experimental infrastructure alters microclimate

160 Experimental structures themselves can alter temperature and other important biotic and abiotic variables in
161 ways that are not generally examined in experimental climate change studies. The importance of controls that
162 mimic a treatment procedure without actually applying the treatment is widely acknowledged in biology (e.g.,
163 Dayton, 1971; Spector, 2001; Johnson & Besselsen, 2002; Quinn & Keough, 2002). Though some experimental
164 climate change studies include treatments with non-functional warming equipment as well as ambient controls,
165 the magnitude and effects of experimental infrastructure alone on climate are rarely interpreted or analyzed.
166 To investigate the magnitude of infrastructure effects, we compared temperature and soil moisture data from
167 five active-warming studies at two sites: Duke Forest and Harvard Forest (Farnsworth *et al.*, 1995; Clark
168 *et al.*, 2014b; Marchin *et al.*, 2015; Pelini *et al.*, 2011)(see Supplemental Materials for model details). These

169 were the only studies in the MC3E database that monitored climate in two types of control plots: structural
170 controls (i.e., ‘shams’ or ‘disturbance controls,’ which contained all the warming infrastructure, such as soil
171 cables (n=1), forced air chambers (n=2), or both (n=2), but with no heat applied) and ambient controls with
172 no infrastructure added. Other studies monitored environmental conditions in only structural controls (n=4)
173 or ambient controls (n=5). We were unable to compare ambient and structural controls for experiments
174 using infrared heating, because no studies in our database included both control types. A separate analysis
175 suggested that there may be infrastructure effects on microclimate for infrared studies in our database (see
176 Supplemental Materials, especially Table S7), and infrastructure effects have been documented in other
177 studies (e.g., shading Kimball *et al.*, 2008).

178 We found that experimental structures altered above-ground and soil temperatures in opposing ways: above-
179 ground temperatures were higher in the structural controls than in ambient controls, whereas soil tempera-
180 tures were lower in structural controls compared with ambient controls (Figure 3a-d). This general pattern
181 was consistent across different temperature models (mean, minimum, and maximum temperatures), although
182 the magnitude varied among seasons, studies, and years (Figure 3a-d, Tables S8-S11). We also found that
183 experimental infrastructure decreased soil moisture relative to ambient conditions across all seasons, studies,
184 and years (Figure 3e, Tables S12, S13).

185 There are several possible reasons for the observed climatic differences between ambient and structural con-
186 trols. Infrastructure materials may shade the plots, reduce airflow, reduce albedo relative to surroundings,
187 or otherwise change the energy balance, particularly in chamber warming (i.e., 4 of the 5 studies included in
188 the above analysis). Specifically, soil temperatures may be cooler in structural controls for forced air studies
189 because the experimental structures block sunlight from hitting the ground surface, causing less radiative
190 heating of the ground in structural controls compared to ambient controls. In addition, above-ground tem-
191 peratures may be warmer in structural controls because the structures radiatively warm the air around them
192 and block wind, inhibiting mixing with air outside of the plot. Structures may also interfere with precipi-
193 tation hitting the ground, thereby reducing local soil moisture and snowpack, with its insulative properties.
194 These effects may be most dramatic in studies that employ chambers, rather than open-air designs, such as
195 infrared heating (Aronson & McNulty, 2009). Finally, for some warming types (e.g., soil cables), structural
196 controls experience increased soil disturbance compared with ambient controls; this may alter water flow and
197 percolation, and introduce conductive material via the cables or posts.

198 To the extent that differences between ambient and structural controls have been reported in previous studies,
199 our findings appear to be consistent. Clark *et al.* (2014b), who used forced air and soil cables with feedback
200 control for warming, state that “control of the air temperature was less precise, in part due to air scooping on
201 windy days.” Marchin *et al.* (2015), who used forced air warming with feedback control, note that structural
202 controls had mean spring air temperatures about 0.5°C or more above ambient temperatures. Peterjohn *et*
203 *al.* (1994), who warmed soil with heating cables and feedback control, reported cooler soil temperatures
204 in structural controls than in ambient controls at shallow soil depths. Similarly, we found the greatest
205 difference in soil temperature between structural and ambient controls in shallow soils (e.g., exp10, in which
206 soil temperature was measured at a depth of 2cm). If addressed, the focus to date has been largely on these
207 abiotic impacts of experimental structures, but structures may also alter herbivory and other biotic conditions
208 (Kennedy, 1995; Moise & Henry, 2010; Wolkovich *et al.*, 2012; Hoeppner & Dukes, 2012).

209 Our analyses suggest that warming experiments that calculate focal response variables relative to ambient
210 controls (e.g., Price & Waser, 1998; Dunne *et al.*, 2003; Cleland *et al.*, 2006; Morin *et al.*, 2010; Marchin *et al.*,
211 2015) may not adequately account for the ways in which infrastructure affects microclimate. Results from
212 studies reporting only structural controls (e.g., Sherry *et al.*, 2007; Hoeppner & Dukes, 2012; Rollinson &
213 Kaye, 2012), should be cautiously applied outside of an experimental context, as—without ambient controls—
214 their inference is technically limited to the environment of the structural controls. Our results suggest that
215 studies aiming to predict or forecast the effects of climate change on organisms and ecosystems would benefit
216 from employing both structural and ambient controls so that they may separate artifacts due to infrastructure
217 from the effects of experimental warming. Increased use of both structural and ambient controls together
218 would also help answer important questions of how infrastructure effects vary across ecosystem types, and
219 warming designs (Box 1).

220 Indirect and feedback effects of climate change manipulations

221 Climate change experiments often seek to manipulate temperature or precipitation separately as well as
222 interactively. Manipulating either of these variables in isolation is notoriously difficult. Treatments involving
223 precipitation additions typically reduce temperatures in climate change manipulations (Sherry *et al.*, 2007;
224 Rollinson & Kaye, 2012; McDaniel *et al.*, 2014b). For example, Sherry *et al.* (2007) observed that a doubling
225 of precipitation reduced mean air temperatures by 0.44°C, on average, during their one-year observation

226 period.

227 In the MC3E database, there are four experiments that manipulated both temperature and precipitation, and
228 provided daily above-ground temperature data (three of these also measured soil temperature). Across these
229 studies, all of which used infrared heating, we found that increasing the amount of added precipitation reduced
230 daily minimum and maximum above-ground temperatures, at rates of 0.01 and 0.02°C, respectively, and soil
231 temperatures, at a rate of 0.01°C for both minimum and maximum temperature, per percent increase in added
232 precipitation (Table S14). Thus, a 50% increase in precipitation would be expected to decrease temperature
233 by 0.5°C. This is likely because increasing soil moisture (an effect of precipitation additions) typically shifts
234 the surface energy balance to favor latent (i.e., evapotranspiration) over sensible energy fluxes, reducing
235 heating of the air overlying the soils. Maintaining target warming levels is a challenge even for independent
236 feedback systems, which vary energy inputs using ongoing temperature measurements, particularly during
237 seasons or years with wetter soils and higher evapotranspiration (Rich *et al.*, 2015).

238 In addition to its effects on temperature, experimental warming often increases vapor pressure deficit and
239 reduces soil water content (e.g., Harte *et al.*, 1995; Sherry *et al.*, 2007; Morin *et al.*, 2010; Pelini *et al.*, 2014;
240 Templer *et al.*, 2016). Of the 15 experiments in the MC3E database, we examined the 12 that continuously
241 measured and reported soil moisture. We included target warming, warming type, and their interaction
242 as predictors (excluding data from plots with precipitation treatments) and accounted for other differences
243 among studies by including a random effect of study (see Supplemental Materials for details). We found
244 that experimental warming reduced soil moisture across all warming types, with substantial variation among
245 experiments (Figure 5, Table S15). The drying effect varied by warming type (-0.80% for infrared versus -
246 0.33% for forced air, per °C of target warming, Table S16). Soil moisture can be difficult to measure, with high
247 spatial and temporal variation (Famiglietti *et al.*, 1999; Teuling & Troch, 2005), but these results highlight
248 that changes in soil moisture often accompany temperature changes in active-warming experiments.

249 Warming and precipitation treatments, and their indirect effects on soil moisture and other abiotic factors,
250 can also alter the biotic environment, which may produce cascading effects. Many studies have found shifts
251 from herbaceous to woody plant communities over time with experimental warming (e.g., Rollinson & Kaye,
252 2012; McDaniel *et al.*, 2014b,a; Harte *et al.*, 2015). These community shifts may affect resource levels, such
253 as moisture, carbon, and nutrient levels in the soil (McDaniel *et al.*, 2014b,a; Harte *et al.*, 2015) and feed
254 back to affect microclimate (Harte *et al.*, 2015).

255 The presence of these feedback effects is both a strength of and a challenge for climate change experiments.
256 They may represent important and ecologically realistic effects that became apparent only with the *in situ* field
257 experiment. Alternatively, they may represent artifacts that are unlikely to occur outside of an experimental
258 context. Quantifying, interpreting, and reporting these non-temperature effects in experiments is critical
259 to distinguish these possibilities and to understand mechanisms underlying observed biological responses to
260 climate change.

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

270 Ecological implications

271 We have highlighted a suite of factors that complicate interpretation of climate change experiments. These
272 indirect effects are likely to have biological implications for many of the responses studied in warming exper-
273 iments (e.g., Figure 5). Interpretation of experimental climate change effects on biological responses may be
274 misleading because the intended climate treatments (i.e., categorical comparisons or target warming levels)
275 are often used as explanatory variables in analyses (Table S1). The interpretation is likely to be altered by
276 using fine-scale, measured climate as explanatory variables. For example, biological responses may be muted
277 (Figure 5b) or exaggerated (Figure 5c) when direct and indirect effects of climate manipulations interact.

278 To investigate the ecological implications of non-target abiotic responses to climate warming, we used a
279 simple case study of plant phenology. We used the MC3E database to test if estimates of the temperature
280 sensitivity of phenology vary when calculated using target warming versus plot-level climate variables. We
281 fit two separate mixed-effects models, which differed in their explanatory variables: one used target warming

and one used measured climate. Both models had budburst day of year as the response variable, and both included random effects of study (which modeled other differences between studies that may have affected phenology), year (nested within study, which modeled differences due to weather variability among years that may have altered phenology), and species (which often vary in their phenology). All random effects were modeled on the intercept only; see Supplemental Materials for details.

We found that phenological sensitivities to temperature estimated from the two modeling approaches varied three-fold. The target warming model estimated temperature sensitivity of budburst to be -1.91 days/°C (95% CI -2.17, -1.86; Table S17, solid black line in Figure 6), whereas the measured climate model estimated temperature sensitivity of budburst to be -6.00 days/°C (95% CI: -6.74, -5.26; Table S17). Further, all measured climate models with both temperature and moisture had improved model fit compared to the target warming model (Table S18). 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 are important drivers of budburst timing (Tables S17, S18). In addition, the measured climate model estimated a significant effect of soil moisture on budburst of -1.51 days/% VWC (95% CI: -1.76, -1.26; Table S17, Figure 6). This negative effect is expected, if reducing moisture delays budburst (Table S17, 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, plot-level warming often does not reach target levels (Figure 2), producing a muted effect of temperature in models using target warming. Second, experimental warming's dual effects of decreasing soil moisture and increasing temperature impact budburst in contrasting ways. Decreasing soil moisture has a delaying effect on budburst phenology, opposing the advancing effect of rising temperatures (Figure 5b); thus the effect of temperature is underestimated when moisture is not included in the model. This example shows how the common method of using target warming alone, or even measured temperature alone as done in previous analyses of the particular experiments included here (exp01, exp03, exp04, exp10, Clark *et al.*, 2014a,b; Polgar *et al.*, 2014; Marchin *et al.*, 2015), to understand biological responses may yield inaccurate estimates of temperature sensitivity in warming experiments. In this case, the underestimation may be substantial enough to account for previously described discrepancies between phenological responses to warming in observational versus experimental studies (Wolkovich *et al.*, 2012; Polgar *et al.*, 2014), though further investigation is required.

311 Accounting for both direct and indirect effects of warming is critical for accurate interpretation of the conse-
312 quences of climate change (Kharouba *et al.*, 2015). Of particular importance is the extent to which abiotic
313 and biotic effects are realistic forecasts of future shifts that are likely to occur with climate change, or due to
314 artifacts that are unlikely to occur outside of experimental systems (Hurlbert, 1984; Moise & Henry, 2010;
315 Diamond *et al.*, 2013). For many important climatic and ecological metrics, experimental findings of abiotic
316 and biotic effects appear to be consistent with observations. Altered above-ground daily temperature range
317 (i.e., temperature minima changing more than maxima, Table S3) with experimental warming is consistent
318 with observed changes in many places. Global minimum temperatures increased more rapidly than max-
319 imum temperatures from 1950-1980, reducing above-ground daily temperature range (Thorne *et al.*, 2016;
320 Vose *et al.*, 2005). In addition, the acclimation response of leaf respiration to temperature (Aspinwall *et al.*,
321 2016; Reich *et al.*, 2016), responses of soil respiration to warming (Carey *et al.*, 2016), and declines in soil
322 carbon at one site (Harte *et al.*, 2015), also appear to be consistent across experiments and observations.
323 These cases suggest that many responses observed in climate change experiments, including indirect effects of
324 treatments, may be accurate harbingers of future biological responses to climate change.

325 In contrast, some responses documented in climate change experiments may not be in line with future
326 climate change—or may be too uncertain for robust prediction, and thus need explicit analyses and cautious
327 interpretation. For example, precipitation forecasts with climate change are uncertain, as are future changes
328 in soil moisture. Soil drying in conjunction with future warming is forecasted in some regions, such as
329 the southwestern United States, mainly because of reductions in precipitation and increased evaporative
330 demand associated with warmer air (Dai, 2013; Seager *et al.*, 2013). The northeastern United States, on
331 the other hand, has been trending wetter over time (Shuman & Burrell, 2017), even though temperatures
332 have warmed. Shifts in soil moisture are likely to vary by region, season, and even soil depth (Seager
333 *et al.*, 2014; Berg *et al.*, 2017). Thus, the soil drying observed in warming experiments may not occur at
334 all sites with future warming. The uncertainty associated with forecasting changes to soil moisture makes
335 replicating future water availability regimes in climate change experiments especially challenging; one way to
336 meet this challenge and make predictions—even given high uncertainty—is to quantify soil moisture effects in
337 climate change experiments. The altered light, wind, and herbivory patterns documented under experimental
338 infrastructure (Kennedy, 1995; Moise & Henry, 2010; Wolkovich *et al.*, 2012; Hoepfner & Dukes, 2012; Clark
339 *et al.*, 2014b) represent other non-temperature effects that may be potential experimental artifacts and are
340 worth quantifying in future analyses to provide improved estimates of temperature sensitivity.

341 An additional challenge in relating experiments to observations is that experimental findings may not scale
342 up in space and time. Short-term responses to climate change frequently differ from long-term responses
343 (Woodward, 1992; Elmendorf *et al.*, 2012; Andresen *et al.*, 2016; Reich *et al.*, 2018). Differences may be, in
344 part, because many experiments typically impose some mean shift in climate, but patterns of climate change
345 are likely to be more variable. Many climate models project complex shifts in precipitation: more intense
346 extreme precipitation events (e.g., heavy downpours), more dry days (i.e., less total precipitation events),
347 or both (Polade *et al.*, 2014). In addition, the small spatial scale of experiments may result in responses
348 that are unlikely to be observed at larger scales (Woodward, 1992; Menke *et al.*, 2014). Experimental plots
349 range in area from 1.5 to 36 square meters (Table S1), which may be too small to encapsulate, for example,
350 the rooting zones of perennial plants (Canadell *et al.*, 1996), or foraging ranges for animals (Menke *et al.*,
351 2014). One approach to overcome these challenges is to conduct larger, longer experiments (Woodward,
352 1992), though this frequently is not logistically possible and does not easily address how to capture potential
353 shifts in climate variability.

354 Conclusions

355 As climate change continues across the globe, ecologists are challenged to not only document impacts, but
356 also make quantitative, robust predictions. Our ability to meet this challenge requires a nuanced mechanistic
357 understanding of how climate directly and indirectly alters biological processes. Climate change experiments,
358 which have been underway for nearly four decades (e.g., Tamaki *et al.*, 1981; Carlson & Bazzaz, 1982; Melillo
359 *et al.*, 2017), provide invaluable information about biological responses to climate change. Yet the full range of
360 changes in environmental conditions imposed by these experiments is rarely presented, and we need a fuller
361 understanding of the variable effects across different warming methodologies. We have compiled the first
362 database of microclimate data from multiple warming experiments and shown how time, space, experimental
363 artifacts, and indirect effects of treatments may complicate simple interpretations of these experimental
364 results. The relative importance of each of these factors is likely to vary across warming designs (Box 1) and
365 myriad attributes of sites, making more studies that measure climate similarly and include full infrastructure
366 controls important for progress. We hope this work provides a foundation for gaining the most knowledge
367 and utility from existing experiments via robust analyses, for designing new experiments (see Box 2), and for
368 improved understanding of biological responses to a changing world.

369 Acknowledgements

370 We are grateful to those who shared their experimental climate data with us, allowing it to be included in
371 the MC3E database. We thank the Radcliffe Institute for Advanced Study at Harvard University, which
372 provided funding for an Exploratory Seminar at which the ideas in this paper were conceived, and we thank
373 three anonymous reviewers. This research was also supported by the National Science Foundation (NSF DBI
374 14-01854 to A.K.E.). Any opinion, findings, and conclusions or recommendations expressed in this material
375 are those of the authors and do not necessarily reflect the views of the National Science Foundation.

³⁷⁶ **Figures**

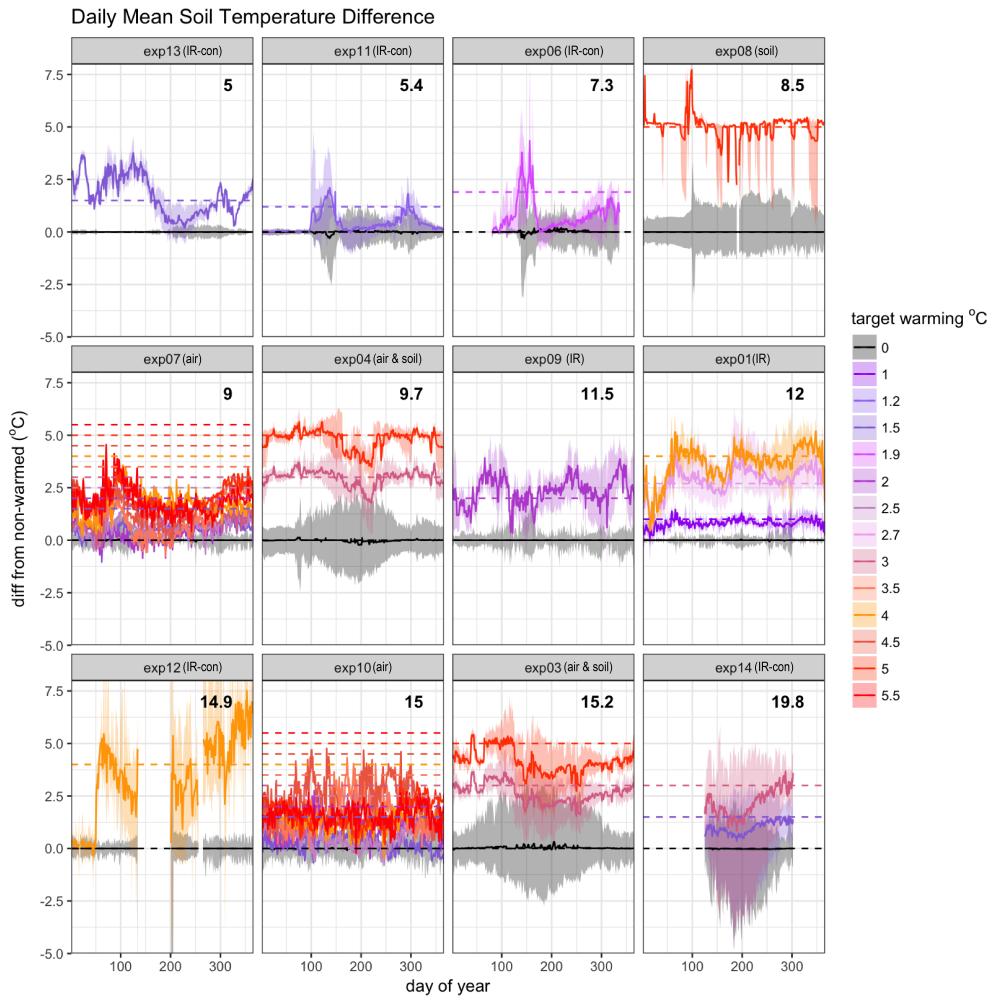


Figure 1: **Deviations in daily observed warming from mean control soil temperature for 12 study sites**, excluding data from plots that manipulated precipitation. We show soil, rather than above-ground, temperature, as this was the most frequently recorded temperature variable in the MC3E database. 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. (Note that the following sites had no explicit target temperature: exp06, exp11, exp12; in exp01, only the highest warming treatment had a target temperature; for these studies and treatments, we used their reported level of warming.) Two sites not shown here did not monitor soil temperature. 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 parentheses next to the site number (IR= infrared, soil= soil cables, air= forced air; con= constant wattage warming control).

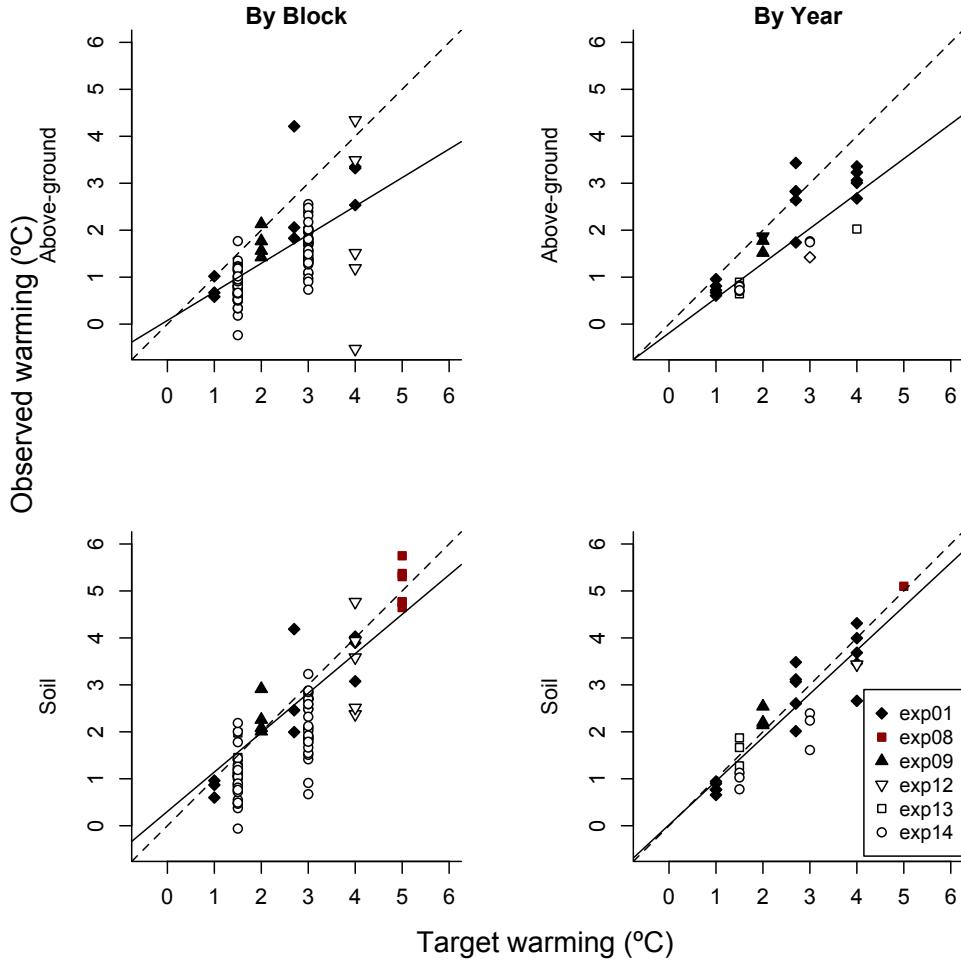


Figure 2: Observed warming over space and time, for above-ground and soil temperatures, excluding data from plots that manipulated precipitation. Above-ground temperature includes air, canopy, and surface temperature. Points represent the difference between treatment and control plots by block (i.e., 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). Black symbols represent studies using infrared; red represents soil warming cables (only exp08); no studies with forced air heating used a blocked design. Open symbols represent constant wattage control and filled symbols represent feedback control. Note that the following studies had no explicit target temperature: exp06, exp11, exp12; for these studies, we used their reported level of warming. For exp01, only the treatment with the greatest warming had a target temperature. See Supplemental Materials (especially Tables S5 and S6) for details.

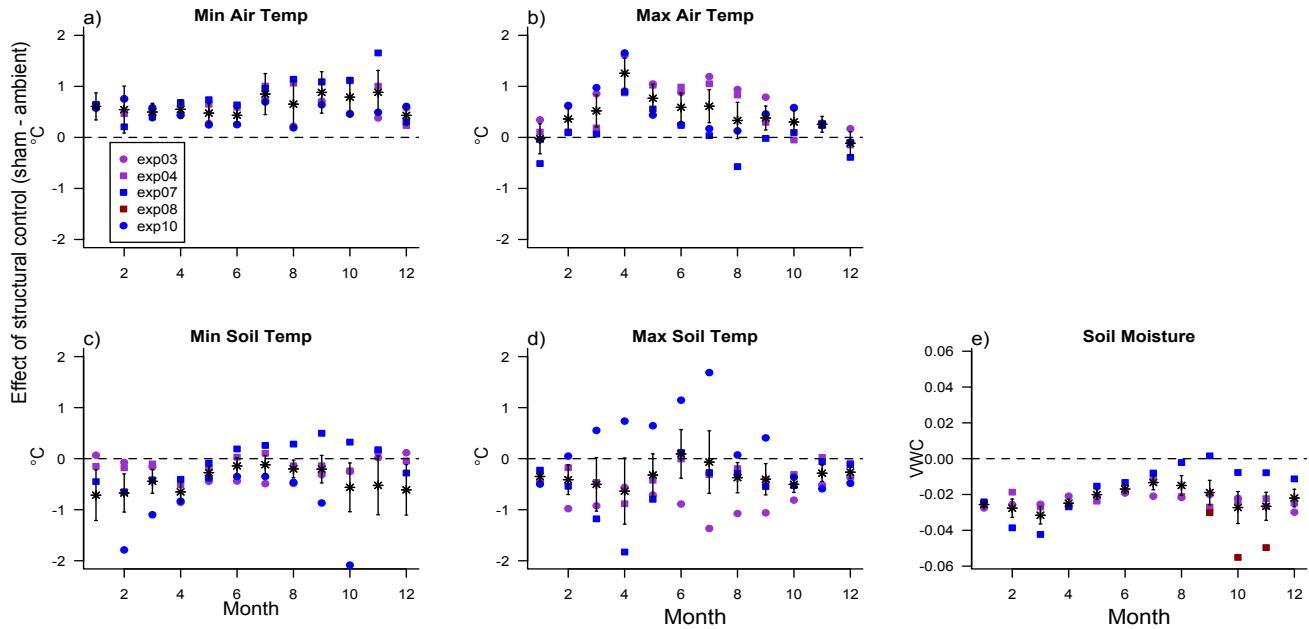


Figure 3: Deviations in measured abiotic variables by month in structural controls compared to ambient controls (i.e., with no control chambers or warming infrastructure in place). Above-ground temperatures (which include includes air, canopy, and surface temperatures) were higher (a,b), whereas soil 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 squares (for the three studies conducted at Harvard Forest in Massachusetts, USA) and circles (the two studies conducted at Duke Forest in North Carolina, USA). Colors vary by heating type: red represents soil warming cables, blue represents forced air; purple represents combined soil warming cables and forced air heating (no studies with infrared heating included both control types). See Supplemental Materials for details (Tables S8-S13).

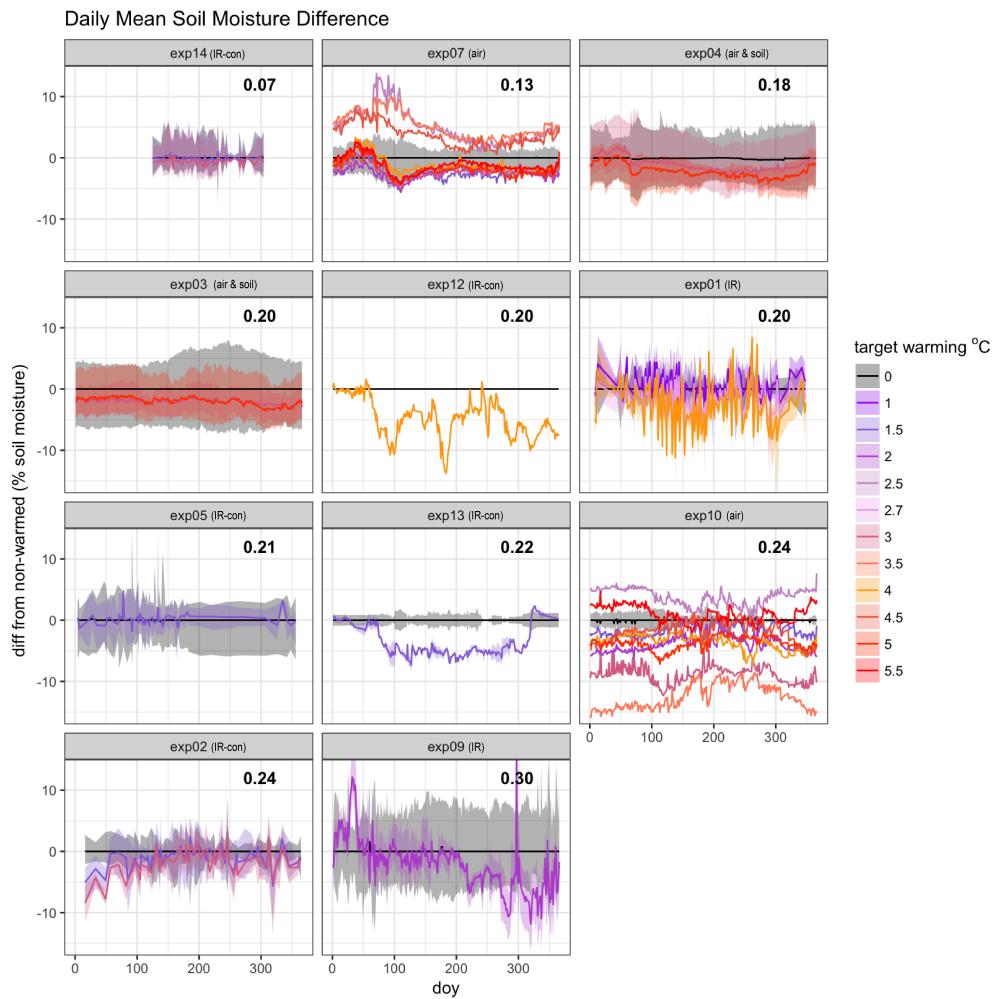


Figure 4: **Deviations in daily observed soil moisture**, shown for the 11 study sites that continuously monitored soil moisture, excluding data from plots that manipulated precipitation. Black lines represent control plots, and colored lines represent warming treatments with various target warming levels (or reported warming, if there was no explicit target temperature). The number of temperature treatment levels vary from one (e.g., exp08, exp11) to nine (exp07 and exp10, which used an unreplicated regression design). 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; the absolute difference between treatment and control plots is shown. Heating type is listed in parentheses next to the site number: IR= infrared, soil= soil cables, air= forced air; con= constant wattage (as opposed to feedback) control.

377 Box 1: Different methods for achieving warming.

378 Active-warming experiments may differ both in the way that they achieve warming (“warming type” in Table
 379 S1), and the way that warming is controlled (“warming control” in Table S1). There are three warming types
 380 used by studies in the MC3E database. These are infrared (n=9), an open-air method in which infrared

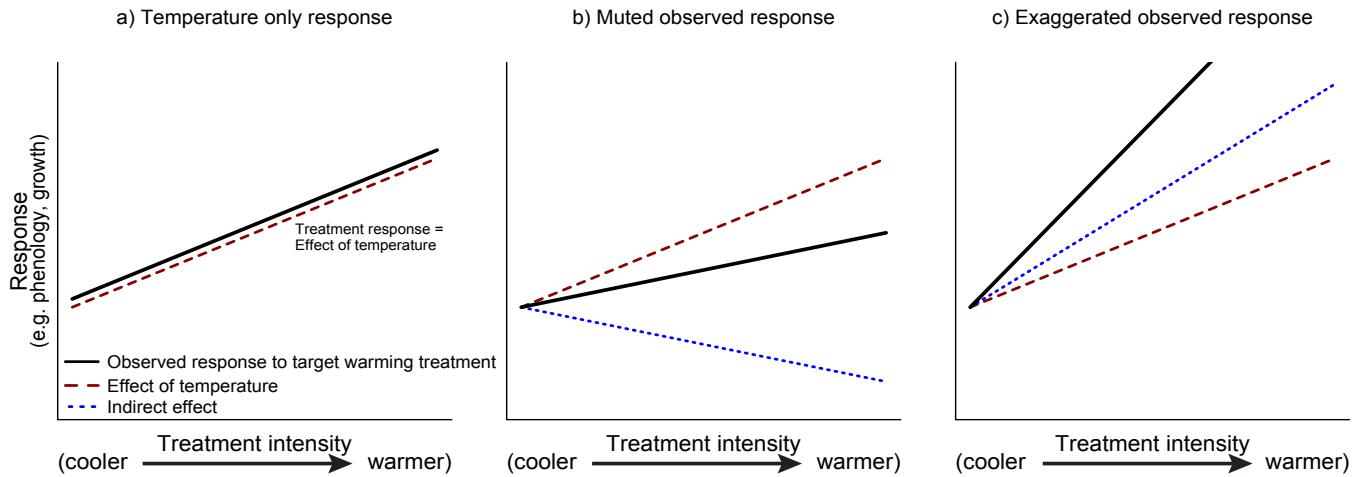


Figure 5: Theoretical biological responses to experimental warming and their interpretation. Direct responses to temperature alone (a) can be easily understood. Complications arise when biological responses are a mix of the direct temperature and indirect non-temperature effects of experimental warming. 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).

heaters are mounted above the ground; forced air ($n=3$), in which air is heated and then pumped through an airflow system into a chamber; and soil warming ($n=1$), a chamber-less method in which soil is heated with buried electric resistance cable. Two additional studies in the database used combined forced air and soil warming in chambers. Warming is maintained by either constant energy output, in which an unvarying energy wattage is used, or by feedback control, in which energy outputs are linked to a thermometer and varied depending on the measured temperature in plots, in order to maintain consistent warming levels.

In this paper, we describe complications and non-temperature effects associated with active warming experiments, across divergent warming methodologies. Some of the non-temperature effects described may be more likely to occur with particular methods than others. Alterations to airflow, for example, may be most dramatic with methods employing chambers. Plot shading and precipitation interference are likely to occur in chamber and infrared techniques, which both involve above-ground infrastructure, and less likely in methods that only warm from the soil. The biological impacts of such effects may be further enhanced or muted based on site characteristics (e.g., if a site is already heavily shaded, impacts from infrastructure shading may be lower).

395 Table 1 highlights that there may be differences in non-temperature effects across these different warming
396 methodologies. In the MC3E database, sample sizes within each warming and control type were quite
397 low, so we were unable to statistically distinguish differences in non-temperature effects across the different
398 methods. We expect the list of non-temperature effects in Table 1 is not exhaustive, but represents what
399 we can document here or has been documented previously. We recommend additional detailed studies of
400 these, and other, effects across warming designs. This will allow future researchers to more fully evaluate
401 the challenges and opportunities of each method, and select an experimental approach well-suited to their
402 particularly research focus.

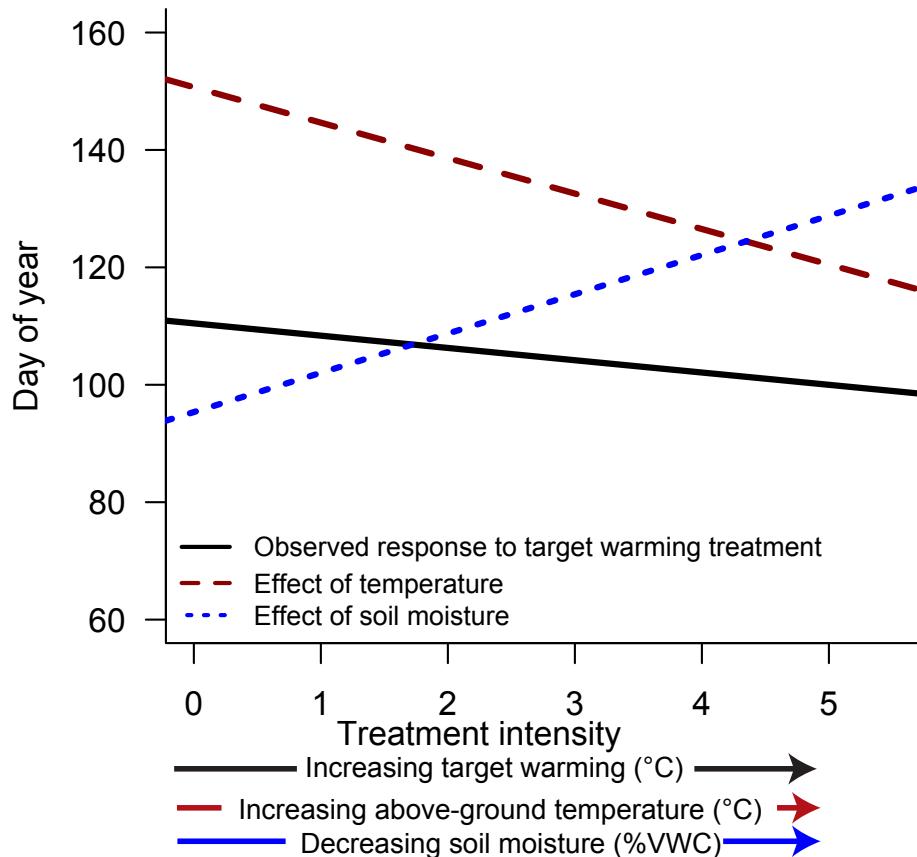


Figure 6: Observed response of budburst day of year to experimental climate change is an example of a muted response: the observed response to increasing treatment intensity (i.e., the coefficient of a model fit with only target [or reported, if there was no explicit target] temperature as the explanatory 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. Analysis includes all studies that monitored budburst and measured soil moisture and above-ground temperature (exp01, exp03, exp04, exp07, exp10); structural control data were used for this analysis (ambient controls were excluded from those studies that contained both). See Supplemental Materials, especially Tables S17 & S18, for additional details.



Soil cable warming experiment at Harvard Forest, Petersham, Massachusetts, USA (Melillo *et al.* 2017). Photograph credit: Yann Vitasse.



Forced air chamber warming experiment at Duke Forest, Hillsborough, North Carolina, USA (exp10, Pelini *et al.*, 2011). Photo credit: Aaron Ellison.



Infrared warming experiment in Pennsylvania, USA (exp09, Rollinson *et al.*, 2012). Photo credit: Christy Rollinson.

Figure 7

Table 1: **Summary of measured warming and documented non-temperature effects, by warming technique**, for studies included in the MC3E database. Summaries of the target warming treatments ($^{\circ}\text{C}$) and measured warming for above-ground temperature, and soil temperature are given. Measured warming is standardized per degree of target warming, and is shown here for warming treatments only (precipitation treatments are excluded). Thus, measured warming is the difference between mean annual temperature (MAT) of control plots and MAT of each treatment level within year (and block, if applicable) of a study. Mean difference (with standard error) and the range of differences in warming are shown, across all years (and blocks, if applicable). n is the number of studies of each warming technique and control type combination in the MC3E database. ‘Soil drying’ was observed in all types with MC3E data, and has been reported for most types in previously published case studies. ‘Other non-temperature effects’ have been observed in many individual studies; the MC3E database does not have data to quantify these other effects.

warming type	warming control	study numbers	target (min-max)	aboveground mean (se)	aboveground range	soil mean (se)	soil range	n	soil drying	other nontemperature effects
infrared	constant	5,6,11-14	2.2 (1.2-4)	0.52 (0.03)	-5.22-2.16	0.68 (0.02)	-0.07-1.56	6	Fig. 4, Tab. S15-16, Kimball et al. 2005	+shading (Kimball et. al. 2005)
infrared	feedback	1,2,9	2.5 (1-4)	0.84 (0.05)	-0.19-1.86	0.96 (0.05)	-0.03-1.85	3	Fig.4, Tables S15-S16, Kimball et al. 2005, Sherry et al. 2007	+freeze-thaw cycles (McDaniel et al. 2013) +shading (Kimball et al. 2005) VPD (Morin et al. 2010)
forced air	feedback	7,10,15	3.5 (1.5-5.5)	0.95 (0.02)	0.43-1.73	0.36 (0.02)	-0.42-1.06	3	Fig.3-4 Norby et al. 1997	+VPD (Norby et al. 1997) air flow (Norby et al. 1997)
soil cables	feedback	8	5			1.01 (0.02)	0.9-1.08	1	Fig.3-4, Peterjohn et al. 1993	+CO ₂ flux (Peterjohn et al. 1993) +N mineralization (Peterjohn et al. 1993)
force air, soil cables	feedback	3,4	4 (3-5)	0.49 (0.06)	-0.02-0.94	0.75 (0.07)	0.01-1.05	2	Fig.3-4	air flow (Clark et al. 2013)

403 Box 2: Recommendations for future climate change experiments

- 404 1. *Collect and analyze plot-level climate data.* This includes analyzing and interpreting minimum and
405 maximum values, as well as variance and critical thresholds (e.g., the number and duration of freeze-thaw
406 events and accumulated chilling hours, McDaniel *et al.*, 2014b; Vasseur *et al.*, 2014). We suggest saving
407 the raw data from data loggers (often collected at hourly or higher resolution) to allow quantification
408 of variance (and other summaries) at different temporal resolutions. In assessing which frequency of
409 measurements is most appropriate for analyses (e.g., hourly, twice daily), it is critical to consider the
410 chronobiology of the event and organisms of interest. For ants, this might mean that temperatures be
411 monitored every minute (Helm & Shavit, 2017); for bacteria, even more frequently.
- 412 2. *Analyze measured climate variables rather than targets.* There can be substantial variation in the effects
413 of warming and precipitation treatments among plots and across time (Figure 2). Analyzing measured
414 climate will allow much more in-depth understanding of the drivers and biological effects of variation
415 in temperature and moisture.
- 416 3. *Publish high quality, usable data and metadata.* Given that climate manipulations are logistically
417 challenging and expensive (Aronson & McNulty, 2009), and that they often produce a large volume of
418 fine-scale climate data, good curation and data sharing will ensure wider use and deeper understanding
419 of these valuable data. When studying biological implications of a global challenge as large as climate
420 change, progress will come from designing and reporting experiments in ways that facilitate an eventual
421 global data set. Researchers should also be explicit in their warming design (e.g., infrared heating
422 with feedback control or forced air heating with constant energy inputs) to aid future analyses of the
423 performance of different designs, across sites and over time (Box 1).
- 424 4. *Include both structural and ambient controls* and collect, use, and report microclimate and biological
425 data within them. Fewer than half of the studies in our MC3E database reported microclimate data
426 from these two control types (6 out of 15 studies); however, all experiments that did include both
427 control types showed significant effects of infrastructure (Figure 3).
- 428 5. *Design relevant manipulations* by consulting observational records and forecasts, including seasonal
429 and annual variation in projected warming. When it is not possible or desirable to match anticipated
430 changes in climate, studies should report how imposed treatments compare to projected changes and

431 past observations (e.g., Hoover *et al.*, 2014; Zhu *et al.*, 2016). In addition, if continuous treatments are
432 not applied throughout the study, we recommend reporting the seasonality and timing of treatments
433 and monitoring the climate throughout the year.

- 434 6. *Maximize the duration of climate change experiments* by running some experiments for as long as
435 possible, since the magnitude of climate change treatments can vary considerably among years (Figure
436 2). In addition, long-term responses of individuals and populations can differ from transient responses
437 (Saleska *et al.*, 2002; Franklin, 1989; Giasson *et al.*, 2013; Harte *et al.*, 2015). We were able to acquire
438 data extending for ≥ 5 years for only one study in the MC3E database (exp01), restricting our ability
439 to investigate the effect of study length on experimental climate change.
- 440 7. *Conduct syntheses across studies.* As more detailed data are published from experimental climate
441 change studies in divergent ecosystems and warming types, meta-analyses will advance our understand-
442 ing of the ways that warming affects microclimate and biotic interactions. For example, it would be
443 useful to compare microclimate data among studies using infrared warming applied with constant en-
444 ergy inputs versus infrared warming that varies energy inputs based on measured temperatures (Box
445 1).

446 References

- 447 Andresen, L.C., Müller, C., de Dato, G., Dukes, J.S., Emmett, B.A., Estiarte, M., Jentsch, A., Kröel-Dulay,
448 G., Lüscher, A., Niu, S. *et al.* (2016). Shifting impacts of climate change: long-term patterns of plant
449 response to elevated co₂, drought, and warming across ecosystems. In: *Advances in ecological research*.
450 Elsevier, vol. 55, pp. 437–473.
- 451 Aronson, E.L. & McNulty, S.G. (2009). Appropriate experimental ecosystem warming methods by ecosystem,
452 objective, and practicality. *Agricultural and Forest Meteorology*, 149, 1791–1799.
- 453 Aspinwall, M.J., Drake, J.E., Campany, C., Vårhammar, A., Ghannoum, O., Tissue, D.T., Reich, P.B. &
454 Tjoelker, M.G. (2016). Convergent acclimation of leaf photosynthesis and respiration to prevailing ambient
455 temperatures under current and warmer climates in eucalyptus tereticornis. *New Phytologist*, 212, 354–367.
- 456 Berg, A., Sheffield, J. & Milly, P.C. (2017). Divergent surface and total soil moisture projections under global
457 warming. *Geophysical Research Letters*, 44, 236–244.

- 458 Bokhorst, S., Huiskes, A., Aerts, R., Convey, P., Cooper, E.J., Dalen, L., Erschbamer, B., Gudmundsson, J.,
459 Hofgaard, A., Hollister, R.D. *et al.* (2013). Variable temperature effects of open top chambers at polar and
460 alpine sites explained by irradiance and snow depth. *Global Change Biology*, 19, 64–74.
- 461 Canadell, J., Jackson, R., Ehleringer, J., Mooney, H., Sala, O. & Schulze, E.D. (1996). Maximum rooting
462 depth of vegetation types at the global scale. *Oecologia*, 108, 583–595.
- 463 Carey, J.C., Tang, J., Templer, P.H., Kroeger, K.D., Crowther, T.W., Burton, A.J., Dukes, J.S., Emmett,
464 B., Frey, S.D., Heskell, M.A. *et al.* (2016). Temperature response of soil respiration largely unaltered with
465 experimental warming. *Proceedings of the National Academy of Sciences*, 113, 13797–13802.
- 466 Carlson, R.W. & Bazzaz, F.A. (1982). Photosynthetic and growth response to fumigation with so₂ at elevated
467 co₂ for c3 and c4 plants. *Oecologia*, 54, 50–54.
- 468 Chuine, I., Bonhomme, M., Legave, J.M., García de Cortázar-Atauri, I., Charrier, G., Lacointe, A. & Améglio,
469 T. (2016). Can phenological models predict tree phenology accurately in the future? the unrevealed hurdle
470 of endodormancy break. *Global Change Biology*, 22, 3444–3460.
- 471 Chuine, I., Morin, X., Sonié, L., Collin, C., Fabreguettes, J., Degueldre, D., Salager, J.L. & Roy, J. (2012).
472 Climate change might increase the invasion potential of the alien c4 grass setaria parviflora (poaceae) in
473 the mediterranean basin. *Diversity and Distributions*, 18, 661–672.
- 474 Clark, J.S., Melillo, J., Mohan, J. & Salk, C. (2014a). The seasonal timing of warming that controls onset of
475 the growing season. *Global Change Biology*, 20, 1136–1145.
- 476 Clark, J.S., Salk, C., Melillo, J. & Mohan, J. (2014b). Tree phenology responses to winter chilling, spring
477 warming, at north and south range limits. *Functional Ecology*, 28, 1344–1355.
- 478 Cleland, E.E., Chiariello, N.R., Loarie, S.R., Mooney, H.A. & Field, C.B. (2006). Diverse responses of
479 phenology to global changes in a grassland ecosystem. *Proceedings of the National Academy of Sciences of
480 the United States of America*, 103, 13740–13744.
- 481 Cox, P.M., Betts, R.A., Jones, C.D., Spall, S.A. & Totterdell, I.J. (2000). Acceleration of global warming
482 due to carbon-cycle feedbacks in a coupled climate model. *Nature*, 408, 184–187.
- 483 Dai, A. (2013). Increasing drought under global warming in observations and models. *Nature Climate Change*,
484 3, 52–58.

- 485 Dayton, P.K. (1971). Competition, disturbance, and community organization: the provision and subsequent
486 utilization of space in a rocky intertidal community. *Ecological Monographs*, 41, 351–389.
- 487 De Boeck, H.J., Vicca, S., Roy, J., Nijs, I., Milcu, A., Kreyling, J., Jentsch, A., Chabbi, A., Campioli, M.,
488 Callaghan, T. *et al.* (2015). Global change experiments: challenges and opportunities. *BioScience*, 65,
489 922–931.
- 490 Diamond, S.E., Penick, C.A., Pelini, S.L., Ellison, A.M., Gotelli, N.J., Sanders, N.J. & Dunn, R.R. (2013).
491 Using physiology to predict the responses of ants to climatic warming. *Integrative and comparative biology*,
492 53, 965–974.
- 493 Dukes, J.S. & Mooney, H.A. (1999). Does global change increase the success of biological invaders? *Trends*
494 *in Ecology and Evolution*, 14, 135–139.
- 495 Dunne, J.A., Harte, J. & Taylor, K.J. (2003). Subalpine meadow flowering phenology responses to climate
496 change: integrating experimental and gradient methods. *Ecological Monographs*, 73, 69–86.
- 497 Elmendorf, S.C., Henry, G.H., Hollister, R.D., Björk, R.G., Bjorkman, A.D., Callaghan, T.V., Collier, L.S.,
498 Cooper, E.J., Cornelissen, J.H., Day, T.A. *et al.* (2012). Global assessment of experimental climate warming
499 on tundra vegetation: heterogeneity over space and time. *Ecology letters*, 15, 164–175.
- 500 Essiamah, S. & Eschrich, W. (1986). Water uptake in deciduous trees during winter and the role of conducting
501 tissues in spring reactivation. *IAWA Journal*, 7, 31–38.
- 502 Ettinger, A. & Wolkovich, E. (2018). Microclimate from climate change experiments (MC3E).
503 doi:10.5063/F1QV3JQR.
- 504 Famiglietti, J., Devereaux, J., Laymon, C., Tsegaye, T., Houser, P., Jackson, T., Graham, S., Rodell, M.
505 & Ovehlen, P.V. (1999). Ground-based investigation of soil moisture variability within remote sensing
506 footprints during the southern great plains 1997 (sgp97) hydrology experiment. *Water Resources Research*,
507 35, 1839–1851.
- 508 Farnsworth, E., Nunez-Farfán, J., Careaga, S. & Bazzaz, F. (1995). Phenology and growth of three temperate
509 forest life forms in response to artificial soil warming. *Journal of Ecology*, 83, 967–977.
- 510 Field, C.B., Lobell, D.B., Peters, H.A. & Chiariello, N.R. (2007). Feedbacks of terrestrial ecosystems to
511 climate change. *Annual Review of Environment and Resources*, 32, 1–29.

- 512 Franklin, J.F. (1989). Importance and justification of long-term studies in ecology. In: *Long-term studies in*
513 *ecology*. Springer, pp. 3–19.
- 514 Friedlingstein, P., Meinshausen, M., Arora, V.K., Jones, C.D., Anav, A., Liddicoat, S.K. & Knutti, R. (2014).
515 Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate*, 27, 511–526.
- 516 Giasson, M.A., Ellison, A.M., Bowden, R., Crill, P.M., Davidson, E., Drake, J., Frey, S., Hadley, J., Lavine,
517 M., Melillo, J. *et al.* (2013). Soil respiration in a northeastern US temperate forest: a 22-year synthesis.
518 *Ecosphere*, 4, 1–28.
- 519 Gruner, D.S., Bracken, M.E., Berger, S.A., Eriksson, B.K., Gamfeldt, L., Matthiessen, B., Moorthi, S.,
520 Sommer, U. & Hillebrand, H. (2017). Effects of experimental warming on biodiversity depend on ecosystem
521 type and local species composition. *Oikos*, 126, 8–17.
- 522 Hagedorn, F., Martin, M., Rixen, C., Rusch, S., Bebi, P., Zürcher, A., Siegwolf, R.T., Wipf, S., Escape, C.,
523 Roy, J. *et al.* (2010). Short-term responses of ecosystem carbon fluxes to experimental soil warming at the
524 Swiss alpine treeline. *Biogeochemistry*, 97, 7–19.
- 525 Hampe, A. (2004). Bioclimate envelope models: what they detect and what they hide. *Global Ecology and*
526 *Biogeography*, 13, 469–471.
- 527 Harte, J., Saleska, S.R. & Levy, C. (2015). Convergent ecosystem responses to 23-year ambient and manip-
528 ulated warming link advancing snowmelt and shrub encroachment to transient and long-term climate–soil
529 carbon feedback. *Global Change Biology*, 21, 2349–2356.
- 530 Harte, J., Torn, M.S., Chang, F.R., Feifarek, B., Kinzig, A.P., Shaw, R. & Shen, K. (1995). Global warming
531 and soil microclimate: Results from a meadow-warming experiment. *Ecological Applications*, 5, 132–150.
- 532 Helm, B. & Shavit, A. (2017). *Dissecting and reconstructing time and space for replicable biological research*,
533 New Haven, CT: Yale University Press, pp. 233–249.
- 534 Hobbie, S.E., Shevtsova, A. & Chapin III, F.S. (1999). Plant responses to species removal and experimental
535 warming in Alaskan tussock tundra. *Oikos*, 84, 417–434.
- 536 Hoeppner, S.S. & Dukes, J.S. (2012). Interactive responses of old-field plant growth and composition to
537 warming and precipitation. *Global Change Biology*, 18, 1754–1768.

- 538 Hoover, D.L., Knapp, A.K. & Smith, M.D. (2014). Resistance and resilience of a grassland ecosystem to
539 climate extremes. *Ecology*, 95, 2646–2656.
- 540 Hurlbert, S.H. (1984). Pseudoreplication and the design of ecological field experiments. *Ecological monographs*,
541 54, 187–211.
- 542 Ibanez, I., Clark, J.S., Dietze, M.C., Feeley, K., Hersh, M., LaDeau, S., McBride, A., Welch, N.E. & Wolosin,
543 M.S. (2006). Predicting biodiversity change: Outside the climate envelope, beyond the species-area curve.
544 *Ecology*, 87, 1896–1906.
- 545 Johnson, P.D. & Besselsen, D.G. (2002). Practical aspects of experimental design in animal research. *ILAR Journal*, 43, 202–206.
- 546
- 547 Kearney, M. & Porter, W. (2009). Mechanistic niche modelling: combining physiological and spatial data to
548 predict species' ranges. *Ecology Letters*, 12, 334–350.
- 549 Kennedy, A. (1995). Temperature effects of passive greenhouse apparatus in high-latitude climate change
550 experiments. *Functional Ecology*, 9, 340–350.
- 551 Kharouba, H.M., Vellend, M., Sarfraz, R.M. & Myers, J.H. (2015). The effects of experimental warming on
552 the timing of a plant–insect herbivore interaction. *Journal of Animal Ecology*, 84, 785–796.
- 553 Kimball, B. (2005). Theory and performance of an infrared heater for ecosystem warming. *Global Change
554 Biology*, 11, 2041–2056.
- 555 Kimball, B.A., Conley, M.M., Wang, S., Lin, X., Luo, C., Morgan, J. & Smith, D. (2008). Infrared heater
556 arrays for warming ecosystem field plots. *Global Change Biology*, 14, 309–320.
- 557 Marchin, R.M., Salk, C.F., Hoffmann, W.A. & Dunn, R.R. (2015). Temperature alone does not explain
558 phenological variation of diverse temperate plants under experimental warming. *Global Change Biology*,
559 21, 3138–3151.
- 560 Marion, G., Henry, G., Freckman, D., Johnstone, J., Jones, G., Jones, M., Levesque, E., Molau, U., Møl-
561 gaard, P., Parsons, A. *et al.* (1997). Open-top designs for manipulating field temperature in high-latitude
562 ecosystems. *Global Change Biology*, 3, 20–32.

- 563 McDaniel, M., Kaye, J. & Kaye, M. (2014a). Do “hot moments” become hotter under climate change? soil
564 nitrogen dynamics from a climate manipulation experiment in a post-harvest forest. *Biogeochemistry*, 121,
565 339–354.
- 566 McDaniel, M., Wagner, R., Rollinson, C., Kimball, B., Kaye, M. & Kaye, J. (2014b). Microclimate and eco-
567 logical threshold responses in a warming and wetting experiment following whole tree harvest. *Theoretical
568 and Applied Climatology*, 116, 287–299.
- 569 Melillo, J., Frey, S., DeAngelis, K., Werner, W., Bernard, M., Bowles, F., Pold, G., Knorr, M. & Grandy,
570 A. (2017). Long-term pattern and magnitude of soil carbon feedback to the climate system in a warming
571 world. *Science*, 358, 101–105.
- 572 Menke, S.B., Harte, J. & Dunn, R.R. (2014). Changes in ant community composition caused by 20 years of
573 experimental warming vs. 13 years of natural climate shift. *Ecosphere*, 5, 1–17.
- 574 Milcu, A., Puga-Freitas, R., Ellison, A.M., Blouin, M., Scheu, S., Girin, T., Frechet, G., Rose, L., Scherer-
575 Lorenzen, M., Barot, S. *et al.* (2016). Systematic variability enhances the reproducibility of an ecological
576 study. *bioRxiv(beta)*, p. 080119.
- 577 Moise, E.R. & Henry, H.A. (2010). Like moths to a street lamp: exaggerated animal densities in plot-level
578 global change field experiments. *Oikos*, 119, 791–795.
- 579 Moorcroft, P.R. (2006). How close are we to a predictive science of the biosphere? *Trends in Ecology &
580 Evolution*, 21, 400–407.
- 581 Morin, X., Roy, J., Sonié, L. & Chuine, I. (2010). Changes in leaf phenology of three European oak species
582 in response to experimental climate change. *New Phytologist*, 186, 900–910.
- 583 Ohlemüller, R., Gritti, E.S., Sykes, M.T. & Thomas, C.D. (2006). Towards European climate risk sur-
584 faces: the extent and distribution of analogous and non-analogous climates 1931–2100. *Global Ecology and
585 Biogeography*, 15, 395–405.
- 586 Parmesan, C. (2006). Ecological and evolutionary responses to recent climate change. *Annual Review of
587 Ecology Evolution and Systematics*, 37, 637–669.
- 588 Pearson, R.G. & Dawson, T.P. (2004). Bioclimate envelope models: what they detect and what they hide -
589 response to hampe (2004). *Global Ecology and Biogeography*, 13, 471–473.

- 590 Pelini, S., Diamond, S., Nichols, L., Stuble, K., Ellison, A.M., Sanders, N., Dunn, R. & Gotelli, N. (2014).
591 Geographic differences in effects of experimental warming on ant species diversity and community compo-
592 sition. *Ecosphere*, 5, 1–12.
- 593 Pelini, S.L., Bowles, F.P., Ellison, A.M., Gotelli, N.J., Sanders, N.J. & Dunn, R.R. (2011). Heating up the
594 forest: open-top chamber warming manipulation of arthropod communities at Harvard and Duke Forests.
595 *Methods in Ecology and Evolution*, 2, 534–540.
- 596 Peterjohn, W.T., Melillo, J.M., Bowles, F.P. & Steudler, P.A. (1993). Soil warming and trace gas fluxes:
597 experimental design and preliminary flux results. *Oecologia*, 93, 18–24.
- 598 Polade, S.D., Pierce, D.W., Cayan, D.R., Gershunov, A. & Dettinger, M.D. (2014). The key role of dry days
599 in changing regional climate and precipitation regimes. *Scientific reports*, 4, 4364.
- 600 Polgar, C.A., Primack, R.B., Dukes, J.S., Schaaf, C., Wang, Z. & Hoeppner, S.S. (2014). Tree leaf out response
601 to temperature: comparing field observations, remote sensing, and a warming experiment. *International*
602 *journal of biometeorology*, 58, 1251–1257.
- 603 Price, M.V. & Waser, N.M. (1998). Effects of experimental warming on plant reproductive phenology in a
604 subalpine meadow. *Ecology*, 79, 1261–1271.
- 605 Quinn, G.P. & Keough, M.J. (2002). *Experimental design and data analysis for biologists*. Cambridge
606 University Press.
- 607 Reich, P.B., Hobbie, S.E., Lee, T.D. & Pastore, M.A. (2018). Unexpected reversal of C3 versus C4 grass
608 response to elevated CO₂ during a 20-year field experiment. *Science*, 360, 317–320.
- 609 Reich, P.B., Sendall, K.M., Rice, K., Rich, R.L., Stefanski, A., Hobbie, S.E. & Montgomery, R.A. (2015).
610 Geographic range predicts photosynthetic and growth response to warming in co-occurring tree species.
611 *Nature Climate Change*, 5, 148–152.
- 612 Reich, P.B., Sendall, K.M., Stefanski, A., Wei, X., Rich, R.L. & Montgomery, R.A. (2016). Boreal and
613 temperate trees show strong acclimation of respiration to warming. *Nature*, 531, 633–636.
- 614 Rich, R.L., Stefanski, A., Montgomery, R.A., Hobbie, S.E., Kimball, B.A. & Reich, P.B. (2015). Design
615 and performance of combined infrared canopy and belowground warming in the B4WarmED (boreal forest
616 warming at an ecotone in danger) experiment. *Global change biology*, 21, 2334–2348.

- 617 Rollinson, C.R. & Kaye, M.W. (2012). Experimental warming alters spring phenology of certain plant
618 functional groups in an early successional forest community. *Global Change Biology*, 18, 1108–1116.
- 619 Rollinson, C.R. & Kaye, M.W. (2015). Modeling monthly temperature in mountainous ecoregions: importance
620 of spatial scale for ecological research. *Climate Research*, 64, 99–110.
- 621 Saleska, S.R., Shaw, M.R., Fischer, M.L., Dunne, J.A., Still, C.J., Holman, M.L. & Harte, J. (2002). Plant
622 community composition mediates both large transient decline and predicted long-term recovery of soil
623 carbon under climate warming. *Global Biogeochemical Cycles*, 16, 3–1–3–18.
- 624 Seager, R., Neelin, D., Simpson, I., Liu, H., Henderson, N., Shaw, T., Kushnir, Y., Ting, M. & Cook, B.
625 (2014). Dynamical and thermodynamical causes of large-scale changes in the hydrological cycle over north
626 america in response to global warming. *Journal of Climate*, 27, 7921–7948.
- 627 Seager, R., Ting, M., Li, C., Naik, N., Cook, B., Nakamura, J. & Liu, H. (2013). Projections of declining
628 surface-water availability for the southwestern United States. *Nature Climate Change*, 3, 482.
- 629 Shaver, G.R., Canadell, J., Chapin, F.S., Gurevitch, J., Harte, J., Henry, G., Ineson, P., Jonasson, S., Melillo,
630 J., Pitelka, L. *et al.* (2000). Global warming and terrestrial ecosystems: A conceptual framework for analysis
631 ecosystem responses to global warming will be complex and varied. Ecosystem warming experiments hold
632 great potential for providing insights on ways terrestrial ecosystems will respond to upcoming decades of
633 climate change. Documentation of initial conditions provides the context for understanding and predicting
634 ecosystem responses. *BioScience*, 50, 871–882.
- 635 Sheldon, K.S., Yang, S. & Tewksbury, J.J. (2011). Climate change and community disassembly: impacts of
636 warming on tropical and temperate montane community structure. *Ecology Letters*, 14, 1191–1200.
- 637 Sherry, R.A., Zhou, X., Gu, S., 3rd, J.A.A., Schimel, D.S., Verburg, P.S., Wallace, L.L. & Luo, Y. (2007).
638 Divergence of reproductive phenology under climate warming. *Proceedings of the National Academy of
639 Sciences of the United States of America*, 104, 198–202.
- 640 Shukla, J. & Mintz, Y. (1982). Influence of land-surface evapotranspiration on the earth's climate. *Science*,
641 215, 1498–1501.
- 642 Shuman, B.N. & Burrell, S.A. (2017). Centennial to millennial hydroclimatic fluctuations in the humid
643 northeast United States during the holocene. *Quaternary Research*, 88, 1–11.

- 644 Spector, R. (2001). Progress in the search for ideal drugs. *Pharmacology*, 64, 1–7.
- 645 Stocker, T., Qin, D., Plattner, G., Tignor, M., Allen, S., Boschung, J., Nauels, A., Xia, Y., Bex, B. & Midgley,
646 B. (2013). IPCC 2013. Climate change 2013: The physical science basis. Contribution of Working Group
647 II to the fifth assessment report of the Intergovernmental Panel on Climate Change.
- 648 Swab, R.M., Regan, H.M., Keith, D.A., Regan, T.J. & Ooi, M.K.J. (2012). Niche models tell half the story:
649 spatial context and life-history traits influence species responses to global change. *Journal of Biogeography*,
650 39, 1266–1277.
- 651 Tamaki, G., Weiss, M.A. & Long, G.E. (1981). Evaluation of plant density and temperature in predator-prey
652 interactions in field cages. *Environmental Entomology*, 10, 716–720.
- 653 Templer, P.H., Phillips, N.G., Ellison, A.M. & Pelini, S.L. (2016). Ecosystem warming increases sap flow
654 rates of northern red oak trees. *Ecosphere*, 7.
- 655 Teuling, A.J. & Troch, P.A. (2005). Improved understanding of soil moisture variability dynamics. *Geophysical
656 Research Letters*, 32.
- 657 Thomas, C.D., Cameron, A., Green, R.E., Bakkenes, M., Beaumont, L.J., Collingham, Y.C., Erasmus,
658 B.F.N., de Siqueira, M.F., Grainger, A., Hannah, L., Hughes, L., Huntley, B., van Jaarsveld, A.S., Midgley,
659 G.F., Miles, L., Ortega-Huerta, M.A., Peterson, A.T., Phillips, O.L. & Williams, S.E. (2004). Extinction
660 risk from climate change. *Nature*, 427, 145–148.
- 661 Thorne, P., Donat, M., Dunn, R., Williams, C., Alexander, L., Caesar, J., Durre, I., Harris, I., Hausfather, Z.,
662 Jones, P. *et al.* (2016). Reassessing changes in diurnal temperature range: Intercomparison and evaluation
663 of existing global data set estimates. *Journal of Geophysical Research: Atmospheres*, 121, 5138–5158.
- 664 Thuiller, W. (2004). Patterns and uncertainties of species' range shifts under climate change. *Global Change
665 Biology*, 10, 2020–2027.
- 666 Urban, M.C., Tewksbury, J.J. & Sheldon, K.S. (2012). On a collision course: competition and dispersal
667 differences create no-analogue communities and cause extinctions during climate change. *Proceedings of
668 the Royal Society B-Biological Sciences*, 279, 2072–2080.
- 669 Vasseur, D.A., DeLong, J.P., Gilbert, B., Greig, H.S., Harley, C.D., McCann, K.S., Savage, V., Tunney, T.D.

- 670 & O'Connor, M.I. (2014). Increased temperature variation poses a greater risk to species than climate
671 warming. *Proceedings of the Royal Society of London B: Biological Sciences*, 281, 20132612.
- 672 Vose, R.S., Easterling, D.R. & Gleason, B. (2005). Maximum and minimum temperature trends for the globe:
673 An update through 2004. *Geophysical Research Letters*, 32, e01221.
- 674 Williams, J.W. & Jackson, S.T. (2007). Novel climates, no-analog communities, and ecological surprises.
675 *Frontiers in Ecology and the Environment*, 5, 475–482.
- 676 Williams, J.W., Jackson, S.T. & Kutzbach, J.E. (2007). Projected distributions of novel and disappearing
677 climates by 2100 AD. *Proceedings of the National Academy of Sciences of the United States of America*,
678 104, 5738–5742.
- 679 Wolkovich, E.M., Cook, B.I., Allen, J.M., Crimmins, T.M., Betancourt, J.L., Travers, S.E., Pau, S. *et al.*
680 (2012). Warming experiments underpredict plant phenological responses to climate change. *Nature*, 485,
681 494–497.
- 682 Woodward, F. (1992). Predicting plant responses to global environmental change. *New Phytologist*, 122,
683 239–251.
- 684 Zhu, K., Chiariello, N.R., Tobeck, T., Fukami, T. & Field, C.B. (2016). Nonlinear, interacting responses to
685 climate limit grassland production under global change. *Proceedings of the National Academy of Sciences*,
686 113, 10589–10594.