Ecology as a science began well over a century ago, when it was based primarily on observational studies (McIntosh 1985). Toward the end of the 20th century, this led to a more experimental, reductionist approach, focused on testing hypotheses and developing theories. Controlled manipulative experiments have shaped our understanding of ecological principles considerably over the past few decades. However, although suitable for testing and developing theories, individual studies yield site-specific information that is difficult to extrapolate to provide broad generalizations. Today, emerging globally relevant questions in basic and applied ecology require a reconsideration of what approaches would be best for understanding large-scale patterns and processes. It is no longer sufficient to ask whether global warming can alter plant or animal growth, whether acid rain has impacts on aquatic trophic structure, or whether habitat fragmentation reduces biodiversity. Society requires more than theory; it requires an understanding of “how, how much, and when”. Our ability to extrapolate the results from experimental approaches has been chiefly limited by the scopes and scales of the studies that are practically and logistically feasible. For this reason, there have been shifts in methodologies in ecology and the environmental sciences; for example, we have seen a return to observational studies through the use of technological advances (Collins et al. 2006; Sagarin and Pauchard 2010). Another technique that has become widely used is meta-analysis, a quantitative approach to reviewing, integrating, and summarizing large numbers of independent studies (Gurevitch and Hedges 2001).

During the past two decades, a progressive increase in the use of meta-analysis has advanced most academic disciplines, including ecology and the environmental sciences. The strength of the meta-analytic approach lies in the objective and quantitative statistical methodology that allows the synthesis of large numbers of studies (Gurevitch et al. 1992; Aronqvist and Wooster 1995; Harrison 2011). A few examples of ecological topics that have been ideally suited to the meta-analytic approach include the question of whether amphibian declines are a worldwide phenomenon (Houlahan et al. 2000), the effect of dispersal on species diversity (Cadotte 2006), the response of ecosystem nitrogen (N) cycles to N addition (Lu et al. 2011), the effects of grazing on plant traits (Díaz et al. 2007), the relationship between plant species rich-
ness and productivity (Mittelbach et al. 2001), the effects of controlling plant traits on litter decomposition rates (Cornwell et al. 2008), and grassland biomass responses to carbon dioxide (CO₂) and N enrichment (Lee et al. 2010). While there has been an increasing use of meta-analysis, there is also a growing literature that is critical, or at least cautionary, of this quantitative approach. Indeed, Whittaker (2010) referred to meta-analyses as "mega-mistakes", not necessarily to discourage the use of meta-analysis but to urge a more rigorous approach to its use.

The limitations and criticisms of meta-analysis are well documented (Hillebrand and Cardinale 2010; Harrison 2011). Ultimately, the robustness of any meta-analysis relies on the individual studies selected for inclusion. For this reason, it is important to exclude methodologically weak studies and to ensure that the studies chosen are representative of all the methodologically robust studies that have been conducted on a research topic. Even when a representative sample of studies is obtained, meta-analysis is only successful if the studies can be combined in a meaningful manner. Two critical issues influence the subsequent success (in terms of scientific accuracy and progress) of meta-analyses: (1) effect size metrics, based on the calculation used to standardize the different types and scales of possible data measurements and the power of the effects, and (2) standardization of methodology amongst studies. Although it is easy to imagine using a meta-analysis to test a basic ecological question – for instance, the richness–productivity relationship among plant communities – the specific effect size metrics used can make a difference to the outcome and interpretation of the analysis. For example, Whittaker (2010) reviewed three meta-analyses, all focused on plant species richness–productivity relationships, and found widely divergent results among the analyses. Similarly, Hungate et al. (2009), in an assessment of the outcome of four published meta-analyses gauging the effect of elevated CO₂ on soil carbon (C), also found that each study reached a different conclusion.

One of the great benefits of meta-analysis is its quantitative approach to hypothesis testing and the collaborative nature of data compilation. Data used in meta-analyses are not necessarily published in the primary peer-reviewed literature. Even if a published paper may seem appropriate for inclusion in a meta-analysis, the researcher conducting the meta-analysis often needs to contact the original authors for further information on methodology or may even request the raw data. This engenders a positive spirit of inclusion and often encourages collaboration. In some cases, papers used in meta-analyses are not cited but rather included as supplemental material and therefore are not indexed by ISI Web of Science, likely because the list of references would otherwise be longer than journals usually allow (Kueffer et al. 2011). Meta-analyses provide a basis for important and meaningful results that can also lead to an appropriate concept for the future design of ecological experiments.

We contend that the answers to many questions in ecology and the environmental sciences cannot be addressed by meta-analysis alone. A core limitation in the meta-analysis approach is that details of the original study are beyond the control of those who conduct the meta-analyses; instead, they are imposed at the time each individual study was conducted. For examples of differences in experimental design, see Figure 1 for precipitation treatments and Figure 2 for warming treatments. Issues of scale, both temporal and spatial, including geographic variation, cannot be controlled retrospectively. A logical solution is to use a study framework that reduces these issues before the data are collected. Long-term experiments, such as the Park Grass Experiment in the UK, which began in 1856 (Silvertown et al. 2006), and the US Long Term Ecological Research Network program (Knapp et al. 2012), have for years been recognized as vital to understanding ecological processes. However, important global issues, such as climate change, N deposition, non-native invasive species, habitat fragmentation and degradation, and biodiversity loss, urgently require standardized controlled experiments on wide geographic scales, preferably global in some cases, and ones that range across biomes and latitudes in others. Coordinated distributed experiments (CDEs) – experiments run in parallel by several research groups in multiple locations around the globe – have the advantage of simultaneously addressing global environmental problems and exploring general ecological theory, while offering the inherent precision of controlled experiments.

Multi-site investigations that we believe resemble eco-
Examples of ecological CDEs

Perhaps the first ecological study that we are aware of that could be termed a CDE was a series of ozone (O₃) experiments conducted from 1980 to 1987 as part of the National Crop Loss Assessment Network (Heagle 1989). At each of five locations in the US, open-top chambers were used to expose important agronomic crop species to a realistic range of O₃ treatments. These studies, when complemented by companion experiments, provided the best evidence that low-level O₃ pollution can affect plant growth and plant community composition, and they also uncovered the mechanisms behind species differences in sensitivity (Reich 1987; Heagle 1989).

A second example of an early ecological CDE was a Canadian-led investigation to determine whether plant competition intensity was independent of plant biomass (Reader et al. 1994). The experiment was inspired by two prevailing theories of plant community organization that differed in their prediction regarding the relationship between competition and site productivity. One theory (Grime 1973) predicted that the intensity of competition increases with an increase in productivity, while the other (Newman 1973; Tilman 1988) predicted that competition would remain consistent with increasing site productivity. Through a standard neighbor removal experiment carried out in 12 locations on three continents by 20 research teams, it was found that competition intensity was not related to neighbor biomass, except at a single site where the range in neighbor biomass was greater than all the other sites. This study highlighted the importance of an adequate range in the productivity gradient and the use of a consistent index for the measurement of competition.

There have been other examples of the CDE approach (Table 1). Several research topics have been or are being addressed, including: the effects of biodiversity on ecosystem function, the coordinated study of ecological responses to global-change factors across a broad geographic range, and the relationships between tree diversity and ecosystem function. The Free-Air Carbon dioxide Enrichment (FACE) network did attempt to achieve uniformity, so that synthesis and meta-analyses would be relatively simple to carry out (Reich et al. 2006; Leakey et al. 2009). Nevertheless, networks of experiments have typically been restricted geographically (primarily to the US or Europe) and have not featured consistent experimental designs or data collection methods across sites. In addition, global meta-analyses may not reflect important patterns within specific biomes. The latter can also be said of TERACC (Rustad 2008; now replaced by INTERFACE), an international research network of global-change experiments linking experimentalists, ecosystem modelers, and Earth-system modelers. So far, the main outcome of research networks has been the building of data repositories for data sharing, often leading to more meta-analyses.

### Examples of ecological CDEs

- **Canada**
  - **PRECIPNET** (Baldocchi et al. 2001), a global network of 500 micrometeorological tower sites from 30 regional networks across five continents, has been effective at combining ecosystem CO₂ flux data. Likewise, the US National Ecological Observatory Network (NEON; Keller et al. 2008), although restricted to the confines of a single country, is being designed to specifically include sites from 20 eco-climate domains, and aims to facilitate the coordinated study of ecological responses to global-change factors across a wide geographic range.

- **Europe**
  - **NitroEurope** (Sutton et al. 2007), a network of N manipulation sites across Europe; **ECOCRAFT** (Medlyn et al. 1999), a network of field experiments to examine the effects of elevated CO₂ on European forest trees; and **TreeDivNet**, an international network of forest researchers who study the relationship between tree diversity and ecosystem function. The Free-Air Carbon dioxide Enrichment (FACE) experiments conducted from 1980 to 1987 as part of the National Crop Loss Assessment Network (Heagle 1989). At each of five locations in the US, open-top chambers were used to expose important agronomic crop species to a realistic range of O₃ treatments. These studies, when complemented by companion experiments, provided the best evidence that low-level O₃ pollution can affect plant growth and plant community composition, and they also uncovered the mechanisms behind species differences in sensitivity (Reich 1987; Heagle 1989). A second example of an early ecological CDE was a Canadian-led investigation to determine whether plant competition intensity was independent of plant biomass (Reader et al. 1994). The experiment was inspired by two prevailing theories of plant community organization that differed in their prediction regarding the relationship between competition and site productivity. One theory (Grime 1973) predicted that the intensity of competition increases with an increase in productivity, while the other (Newman 1973; Tilman 1988) predicted that competition would remain consistent with increasing site productivity. Through a standard neighbor removal experiment carried out in 12 locations on three continents by 20 research teams, it was found that competition intensity was not related to neighbor biomass, except at a single site where the range in neighbor biomass was greater than all the other sites. This study highlighted the importance of an adequate range in the productivity gradient and the use of a consistent index for the measurement of competition.

- **China**
  - **TERACC** (Rustad 2008; now replaced by INTERFACE), an international research network of global-change experiments linking experimentalists, ecosystem modelers, and Earth-system modelers. So far, the main outcome of research networks has been the building of data repositories for data sharing, often leading to more meta-analyses.
tem function (BIODEPTH; Hector et al. 1999), global warming (ITEX; Henry and Molau 1997), combinations of warming and rainfall manipulation (CLIMOOR/VULCAN/INCREASE; Beier et al. 2004; Wessel et al. 2004), and nutrient additions (NITREX and EXMAN; Wright and Rasmussen 1998; NutNet, Adler et al. 2011). Other networks, such as LIDET (Gholz et al. 2000), which examined differences in rates of litter decomposition, and LINX I and II (Webster et al. 2003; Mulholland et al. 2008), which used 15N tracers to explore stream N cycling, provided useful comparisons of ecosystem processes among sites but without experimental manipulation of the environment. For all of these CDEs, the emphasis has been on simple, testable questions and standardized methodology. However, in most cases, with the exception of the Nutrient Network (NutNet) and the International Tundra Experiment (ITEX), the geographic range has been limited.

ITEX provides some of the first evidence of the effects of global warming on ecosystems and was the first international effort to understand the potential future impacts of climate change. ITEX is circumpolar, involving 10 countries and 26 sites across the Arctic tundra. Considering that global warming is predicted to impact the higher latitudes first and with the greatest intensity, the experiment was instrumental in gaining an understanding of climate-change effects. All plant species investigated under ITEX responded to warming in terms of either biomass changes, phenology, or both; there was, however, no general pattern in terms of type or magnitude of response over the limited time span of the experiment (Henry and Molau 1997), although the cumulative effects of long-term warming may be greater than expected (Elmendorf et al. 2011). The results suggest that not only are ecological CDEs important for a general understanding of pattern and process, but that long-term, multi-site investigations are needed to understand changes in response and how those changes vary among ecosystems (Knapp et al. 2012).

NutNet was developed to address how grasslands are affected by global changes in eutrophication, mainly through atmospheric nutrient deposition, and large herbivore communities (Adler et al. 2011; Firn et al. 2011; Stokstad 2011). NutNet began partly in response to a frustration among researchers with the lack of any data on terrestrial systems that could be used in meta-analyses (Hillebrand et al. 2007; Gruner et al. 2008). The experimental approach was to study the addition of three nutrient categories (N, phosphorus, and potassium plus micronutrients), and the multiple combinations of all nutrient categories and fencing to exclude herbivores. Currently, 40 sites on five continents are implementing

<table>
<thead>
<tr>
<th>Network meta-experiment</th>
<th>Treatments</th>
<th>Technology</th>
<th>Year(s) active</th>
<th>Number of sites (location)</th>
</tr>
</thead>
<tbody>
<tr>
<td>BIODEPTH</td>
<td>Species diversity</td>
<td>Constructed communities</td>
<td>1996–1999</td>
<td>8 (grasslands – Europe)</td>
</tr>
<tr>
<td>CLIMOOR/VULCAN/INCREASE</td>
<td>Warming, rain exclusion</td>
<td>Retractable curtains</td>
<td>1998–present</td>
<td>6 (heathland/shrubland – Europe)</td>
</tr>
<tr>
<td>Competition Intensity</td>
<td>Neighbor removal</td>
<td>Herbicide</td>
<td>1992</td>
<td>12 (old field/grassland – North America/Europe/Australia)</td>
</tr>
<tr>
<td>GLOWA Jordan River</td>
<td>Rain addition/ exclusion</td>
<td>Irrigation and rain-out shelters</td>
<td>2001–present</td>
<td>4 (shrubland/dwarf shrubland – Israel)</td>
</tr>
<tr>
<td>ITEX</td>
<td>Warming</td>
<td>Open-top chambers</td>
<td>1992–present</td>
<td>26 (Arctic tundra)</td>
</tr>
<tr>
<td>LINX I</td>
<td>Site comparison only</td>
<td>15N tracer addition</td>
<td>1996–2001</td>
<td>16 (streams – US)</td>
</tr>
<tr>
<td>LINX II</td>
<td>Site comparison only</td>
<td>15N tracer addition</td>
<td>2001–2006</td>
<td>8 (9 streams per location – US)</td>
</tr>
<tr>
<td>LIDET</td>
<td>Site comparison only</td>
<td>Decomposition – litter bags</td>
<td>1990–2000</td>
<td>28 (North and Central America)</td>
</tr>
<tr>
<td>NITREX</td>
<td>N addition or removal</td>
<td>Irrigation and rain-out shelters</td>
<td>Established late 1980s</td>
<td>8 (coniferous forest – northwest Europe)</td>
</tr>
<tr>
<td>EXMAN</td>
<td>N quantity and quality</td>
<td>Irrigation and rain-out shelters</td>
<td>Established 1980s</td>
<td>6 (coniferous forest – Europe)</td>
</tr>
<tr>
<td>NutNet</td>
<td>Factorial N, P, K addition/vertebrate herbivore exclosure</td>
<td>Fertilizer application and fencing</td>
<td>2006–present</td>
<td>45 (grasslands – global)</td>
</tr>
</tbody>
</table>
the full experimental protocol, while five more sites are implementing only the factorial nutrient experiment.

### What is an ecological CDE?

Partly on the basis of the above-cited examples, we offer the following defining attributes (1–5) and probable characteristics (6–8) of ecological CDEs.

1. **Hypothesis-driven experimental study**
   
   The research question must be direct, testable, and determined through consensus of the collaborative team. This requires that the problem be condensed to an overarching general question that has potential validity in different geographic locations and in ecosystems with differing species pools, legacies (e.g., natural and anthropogenic disturbances), biodiversity, climates, and so forth. This approach forces the research team to identify general and specific research questions.

2. **Multiple geographic locations**
   
   Global issues require multi-site experiments at a transcontinental spatial scale. The majority of meta-analyses seem to be drawn from studies in North America and Europe, leading to the possibility of a bias in outcomes (Martin et al. 2012). The larger the geographic range a CDE encompasses, the greater the potential for understanding general principles or identifying the underlying causes of different responses between site locations.

3. **Standardized research design**
   
   Every research team at each location must follow the same methodology and protocol (based on collaborative consensus) outlined in the experimental design.

4. **Standardized data and coordinated data management**
   
   Ecologists use a variety of data management methods even when recording identical phenomena. Standardizing input forms and data storage allows for more straightforward data compilation and analysis and is critical for taking advantage of the power of CDEs. Such data standardization also facilitates the generation of metadata and increases the potential for alternative uses of the experimental data.

5. **Intellectual property sharing**
   
   A common agreement on data ownership and intellectual property must be based on a consensus among the research teams, so as to avoid conflicts and to maximize commitment of all the partners to the common approach.

6. **Synchronized data collection**
   
   Research groups should run the experiment simultaneously if the different study sites share similar seasonal climates. Alternatively, studies should run during the same growing season if northern and southern hemispheres are included, to reduce the likelihood of temporal variation in global climate or atmospheric composition affecting research outcomes.

7. **Multiple investigative teams**
   
   Networks that include teams of investigators from many locations and disciplines reduce costs for each individual team while promoting multidisciplinary collaborations and increasing the potential for new insights to develop from the sharing of ideas. Having researchers close to each proposed study location reduces travel costs and individual time commitments.

8. **Low cost and low maintenance**
   
   Low-cost activities will encourage collaboration among researchers around the globe, including those in developing countries with limited finances and in sites that are remote and potentially difficult to access by North American and European researchers. Because it is currently a challenge to attract dedicated funding for global collaborative projects, many ecological CDEs will need to be conducted on existing operating budgets. However, ecologists and environmental scientists must be more vocal about educating policy makers and research agencies about the importance of this research, even though it may be expensive by historical ecological standards.

Ecological CDEs require coordination and uniformity in design and application. Because all the data are collected from the same replicated experimental design, the statistical approach to data interpretation is very different from that used in meta-analyses. Meta-analysis involves complex statistics specific to and designed for this purpose (Hedges and Olkin 1985). Statistical analysis of ecological CDEs relies on arguably simpler, traditional tests, such as analysis of variance, mixed-effects models, and multilevel/hierarchical models (Gelman and Hill 2007).

Perhaps the most challenging aspect of ecological CDEs is developing the networks and funding the experiments. Ultimately, it takes one individual, or a small group of people, to assume a leadership role in recruitment and coordination of the network. At the recruitment stage, it is imperative that the basic research question is well formulated and addresses a critical area of interest. Once the question is established, a starting point for recruitment should include personal contacts and internet searches for researchers that are currently engaged in similar research, ascertained through recent publications. Protocol development will require a critical mass of
invested researchers, and while this can be done “remotely” via e-mail, listservs, and social media, we recommend an organized workshop, which can be coordinated to run during an international conference to maximize attendance. We encourage organizing committees of international ecology conferences to accommodate CDE workshops. A well-designed protocol can serve as a recruitment tool to enlarge the network, especially in identified geographical gaps, and as the basis for subsequent grant proposals. Once a CDE is underway, annual or biennial workshops can serve to strengthen the network and maximize productivity. Few funding agencies support large intercontinental ecological experiments, but we expect that the international demand for suitable funding mechanisms will increase as CDEs become more common. We strongly recommend that national and transnational research agencies and foundations establish and promote international agreements for the funding of participating researchers in CDEs.

A proposal for a CDE to test the effects of drought in grasslands

In grassland ecosystems, water is the major limiting factor; changes in precipitation patterns and increased risk of drought will therefore likely have a major impact on grassland ecosystems. Greenhouse-gas emissions have led to increases in atmospheric CO₂ concentration and mean annual global temperature, causing alterations in mean annual global precipitation (Meehl et al. 2007). A key prediction of altered precipitation is an increased risk of drought (Reichstein et al. 2007), and yet relatively few studies have investigated the effect of potential changes in precipitation on terrestrial ecosystems (Weltzin et al. 2003). Predicted changes in precipitation vary across grasslands globally (increase, no change, decrease) but all models agree that, globally, precipitation will be more variable with more extreme events (e.g., droughts). A global chronic drought experiment is therefore needed. Regions of the world can vary widely in projected trends in precipitation changes (Meehl et al. 2007), and yet relatively few studies have investigated the effect of potential changes in precipitation on terrestrial ecosystems (Weltzin et al. 2003). Predicted changes in precipitation vary across grasslands globally (increase, no change, decrease) but all models agree that, globally, precipitation will be more variable with more extreme events (e.g., droughts). A global chronic drought experiment is therefore needed. Regions of the world can vary widely in projected trends in precipitation changes (Meehl et al. 2007), and yet relatively few studies have investigated the effect of potential changes in precipitation on terrestrial ecosystems (Weltzin et al. 2003). Predicted changes in precipitation vary across grasslands globally (increase, no change, decrease) but all models agree that, globally, precipitation will be more variable with more extreme events (e.g., droughts). A global chronic drought experiment is therefore needed. Regions of the world can vary widely in projected trends in precipitation changes (Meehl et al. 2007), and yet relatively few studies have investigated the effect of potential changes in precipitation on terrestrial ecosystems (Weltzin et al. 2003).
change experiments on three continents were inconsistent and it was not possible to identify broad patterns of grassland response to temperature and precipitation manipulations. One of the likely reasons were the many differences in experimental methodology, including the size of rain-out shelters (structures built to reduce or restrict precipitation from reaching the ground), the design of the rain-out shelters, the magnitude of water additions, and the design of open-top chambers to increase temperature (White et al. 2011).

Most grassland climate-change experiments have been carried out in North America and Europe (Figure 3). This geographic bias limits our ability to determine general global patterns of grassland response to global change. The large variation in design and scale of experiments further limits the potential for comparison and development of general principles (Figures 1 and 2). Two meta-analyses have investigated experimental climate-change effects on plant communities (Rustad et al. 2001; Wu et al. 2011), but the same problems inherent in all meta-analyses apply (explained above). We identify the need for a collaborative effort of researchers from around the globe to tackle what we consider to be the most relevant and pressing question about grassland ecosystem responses to climate change, and to do so in the context of a CDE.

The proposed CDE would ask the general question “are the effects of drought on temperate grassland ecosystems consistent across study sites?” The research design would involve installations of standardized, low-cost, passive rain-out shelters (eg Heisler-White et al. 2009; Figure 4) over grassland ecosystems around the world. Passive rain-out shelters require little maintenance and allow for relatively easy set-up and monitoring. Response variables to measure could include plant abundance and diversity, soil microbial communities, levels of soil C and N, and soil moisture. Of course, other response variables would also likely be measured, including plant traits and potentially soil biota, but a bare minimum of site characteristics (such as soil properties) and response variables would be required for a site to be included in the CDE. Such an approach will facilitate testing for general effects across wide site differences.

Advantages of the CDE approach over the meta-analysis approach are evident when examining the consequences of rainfall manipulations. Unlike ecological CDEs, meta-analyses are not designed to test for within-site differences or the complexity of interaction effects. For example, if plant biomass within half of the rain-out shelters in one site responded negatively and the other half in the same site responded positively, a meta-analysis might interpret the site response as having no overall effect. Because data from an ecological CDE are standardized, it is possible to test for interactive site differences, both within a site and among sites, as well as accounting for covarying climate and habitat variables, assuming these properties are also monitored as a part of the CDE.

A potentially important property that varies between geographic locations is year-to-year variation in annual precipitation. If the coefficient of variation over two decades in one site is greater than that in another site, the rain-out shelter treatment might have less of an effect on the former site as compared with the latter. An ecological CDE will be able to answer this question by comparing different grassland sites with different year-to-year variation in annual precipitation; thus, temporal and spatial scale can be controlled. Standardized control of experiments and control of detail in measurement of response across a wide geographic range enhances the power of the CDE approach over that of meta-analysis.

Conclusions

Anthropogenic activities are having effects at global scales. For many of the environmental issues we face it is no longer possible for a single investigator, or even a team of investigators, to address large-scale ecological questions at the local or regional level. Ecological CDEs that rely on international collaboration are needed to address worldwide problems; these will help us to understand ecological patterns and processes and to develop solutions that will benefit environmental management.

References

Coordinated distributed experiments


