CHAPTER 6

SCALING TERRESTRIAL BIOGEOCHEMICAL PROCESSES:

Contrasting Intact and Model Experimental Systems

MARK A. BRADFORD AND JAMES F. REYNOLDS

6.1 INTRODUCTION

Planet Earth is undergoing enormous change, clearly discernable on time scales of decades to centuries. This is largely a result of human activities, especially the emission of greenhouse gases and pollutants, the clearing of global forests, the urbanization of agricultural cropland, and many other extensive modifications in the land surface. Significantly, these rates of change are unprecedented in the Earth's geological history (Committee on Global Change 1988), and one of the great scientific challenges of the 21st century is to forecast future behaviors of global ecosystems under the constant pressure of human insults (Clark et al. 2001). This requires that we better understand feedbacks and interactions of the major patterns and processes of the key components of planet Earth: the atmosphere, oceans, freshwater, rocks, soils, and biosphere. In an attempt to meet this challenge, an interdisciplinary approach to studying systems dynamics on a planetary-scale has emerged, known as Earth System Science (ESS) (Schellnhuber 1999, Lawton 2001).

Biogeochemical processes of terrestrial ecosystems are at the core of ESS research (Schellnhuber 1999). Hence, considerable effort has been invested towards understanding the relative importance of biotic and abiotic regulators and controllers. Of special interest are how natural and human-induced perturbations may affect the rates and directions of biogeochemical processes in terrestrial ecosystems, especially in terms of potential feedbacks to climate systems (Walker and Steffen 1996, Pielke 2002). Given the complexity of global systems and their many interconnections, one of the main scientific challenges of ESS is to document change, diagnose underlying causes, and develop plausible projections of how natural variability and human actions may affect global biogeochemical cycles in the future. With regard to the latter, once we have the requisite quantitative understanding of process rates, as well as a detailed understanding of key regulatory

mechanisms, the goal is to extrapolate findings obtained at one temporal and spatial scale to another.

Typically extrapolations for terrestrial ecosystems are done using mathematical models (see Goudriaan et al. 1999, Prinn et al. 1999, Wu and Li, Chapters 1 and 2). However, in this chapter we focus on extrapolations via empirical experimentation: we discuss experimental designs that inform about process rates and regulatory factors at spatial and temporal scales *greater* than the one on which the experiment is conducted. Paradoxically, while most experiments are in fact *intended* to further understanding and knowledge at scales *beyond* the ones at which they are actually being conducted (e.g., 1 m^2 plots or a forested watershed), most fail to incorporate the spatial and temporal scale considerations necessary to justify such an extrapolation (Gardner et al. 2001). A number of issues are germane to this discussion, including the specific characteristics of the factors under investigation, the importance of nonlinear responses, the type of treatment imposed (e.g., step vs. gradual), and whether the goal is spatial or temporal extrapolation.

We discuss experiments conducted using two general types of systems: intact systems and model systems. Our objective is to compare and contrast these approaches in the context of their potential for contributing to our predictive understanding of process rates and their regulators in terrestrial biogeochemical cycles. In this chapter we will show that: (1) *intact ecosystem experiments* can provide process rates, mechanistic understanding and absolute/relative treatment effects suitable for direct extrapolation, but rarely do; and (2) *model ecosystem experiments* can provide the sign (positive or negative) of treatment effects and insights into their mechanistic basis. However, data obtained on process rates and absolute/relative treatment effects are not suitable for extrapolation. We concur with Gardner et al. (2001) that there is a need for much greater "scale awareness" in ecology, especially with regard to the role of experimental design and execution. Our primary objective is to raise awareness of the importance of spatial and temporal scale considerations in the design and interpretation of experiments, so that findings at the scale of an experimental plot and duration may be extrapolated with *known* confidence.

6.2 DEFINITIONS

For the purposes of our discussion, empirical experimental systems are grouped into two classes: (i) intact and (ii) model. Prior to embarking on our discussion it is important to note that although neither mathematical modeling nor observation are explicitly discussed here, we recognize the critical importance of these approaches, which are adequately covered elsewhere (e.g., Goudriaan et al. 1999, Gardner et al. 2001). Also, much discussion on scaling in experimental ecology is semantic and, albeit this is an important debate, it is beyond the scope of our discussion. The literature is replete with terminology that is often contradictory and/or case-specific; hence, we cite established terminology where possible and provide definitions when introducing context-specific terms.

6.2.1 Intact Ecosystem Experiments

Intact ecosystem experiments (IEEs) utilize plots where the established ecosystem of interest, e.g., a forest, agricultural field or desert, is sampled "intact" by the experimenter. Within the ecosystem of interest, a defined area is selected (arbitrarily, pragmatically or using a statistical rationale) and one or more factors are manipulated or perturbed across replicate plots *in situ*. Treatment effects are quantified by way of comparison with non-manipulated "control" or "reference" plots. Examples of IEEs are the exclusion of exotic, mammalian herbivores using fenced areas across 30 locations (Wardle et al. 2001); surrounding six 30 m diameter forest plots with free air CO_2 enrichment (FACE) rings and exposing half to ambient and half to elevated atmospheric CO_2 concentrations (DeLucia et al. 1999); and establishing ten alpine tundra plots and adding nitrogen fertilizer to five (Neff et al. 2002).

6.2.2 Model Ecosystem Experiments

In contrast, model ecosystem experiments (MEEs) are conducted either within the field or laboratory. We define MEEs as those where the experimental system is either wholly or partially "constructed" and one or more treatments are applied. Well-known field examples are the biodiversity manipulations at Cedar Creek (Tilman et al. 1996) and the pan-European Biodepth study (Hector et al. 1999), where plant communities of varying species richness and composition were constructed. Laboratory-based experimental systems span ranges of biotic complexity from simple Gaussian style systems (e.g., McGrady-Steed et al. 1997), through individual plants in pots of soil grown in controlled environment chambers (e.g., Fernández and Reynolds 2000), to multi-species, multi-trophic systems constructed as analogues of intact systems (e.g., Lawton 1996). A vegetated soil monolith placed within the laboratory would, although an intact unit, be defined as an MEE because its climate is constructed.

6.3 MEASURING TREATMENT EFFECTS USING INTACT ECOSYSTEM EXPERIMENTS

6.3.1 Plot Size and Spatial Extrapolation

We define *internal regulators* as factors that affect some ecosystem process of interest and which, in contrast to *external regulators*, vary within the system. Examples of each are soil fertility and air temperature, respectively. The key to extrapolating results from experimental plots to larger spatial scales is an understanding, or at least appreciation, of which internal factors regulate the process under investigation (e.g., how soil nitrogen affects plant growth; how soil water affects leaf photosynthesis) and how these factors are spatially distributed. To illustrate this point, we use an example whereby we assume that the values of the factors are invariant over time and exert their control internally on the process of

interest. If we have an understanding of what internal regulatory factors are important – and their spatial distribution in the environment – we can opt to select the size of our plots to capture the spatial heterogeneity of these regulatory factors (i) within a single plot (e.g., Figure 6.1a), or (ii) through use of multiple plots, each of which is placed within a spatially homogenous patch of the regulatory factors (e.g., Figure 6.1b).
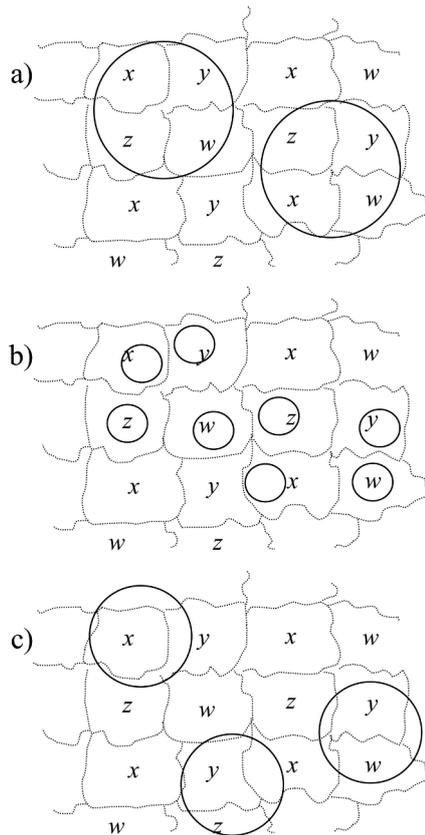

Figure 6.1. Selection of plot sizes (open circles) to capture the spatial heterogeneity of a regulatory factor and, thus, permit spatial extrapolation of the measured process from the plot to the landscape. Letters in italics denote a fixed level of the regulatory factor. There are four levels, w-z, and the spatial distribution of the regulatory factor is repeated every two homogenous areas. In (a) and (b) plot sizes permit extrapolation to the landscape. In (a) the plot size captures the heterogeneity within a single plot. In (b) four plots are required to capture the same heterogeneity but the advantage is that the collected data can be used to construct the shape of the relationship between the rates of the process and the levels of the regulatory factor. Plots can be placed anywhere within the homogenous area of a single value of the regulatory factor. In (c) the plots are inappropriately sized to capture the spatial heterogeneity and so the results from the plots cannot be extrapolated to the landscape. In the scenario depicted in (c) the plots could be positioned to capture an equal proportion of w-z within them but in reality the heterogeneity is unlikely to be so uniformly distributed.

   To capture the spatial heterogeneity of the regulatory factors within our plots we first need to know over what "target" area we plan to extrapolate our results and then measure the spatial heterogeneity of the regulatory factors across this landscape. The advantages of capturing the spatial heterogeneity within our plot size (Figure 6.1a) are: (i) measured process rates can be directly extrapolated across the landscape of interest; (ii) there is an inherent assumption that large plot sizes inform more about how a system would naturally respond to a perturbation than smaller plot sizes; and (iii) as plots increase in size within a fixed spatial area the variation between replicate plots decreases. The disadvantage is that we do not generate data necessary to construct the shape of the relationship between process rates and

regulatory factors. This information is required to parameterize predictive models given that many processes respond nonlinearly to variation in regulatory factors (Figure 6.2a-c). From the relationships depicted in Figure 6.2a-c it is clear that a single mean value of the regulatory factor (as would be obtained by "capturing" the spatial heterogeneity within single plots) cannot be used to determine the shape of the response of the process of interest across varying regulator values (Figure 6.2d).
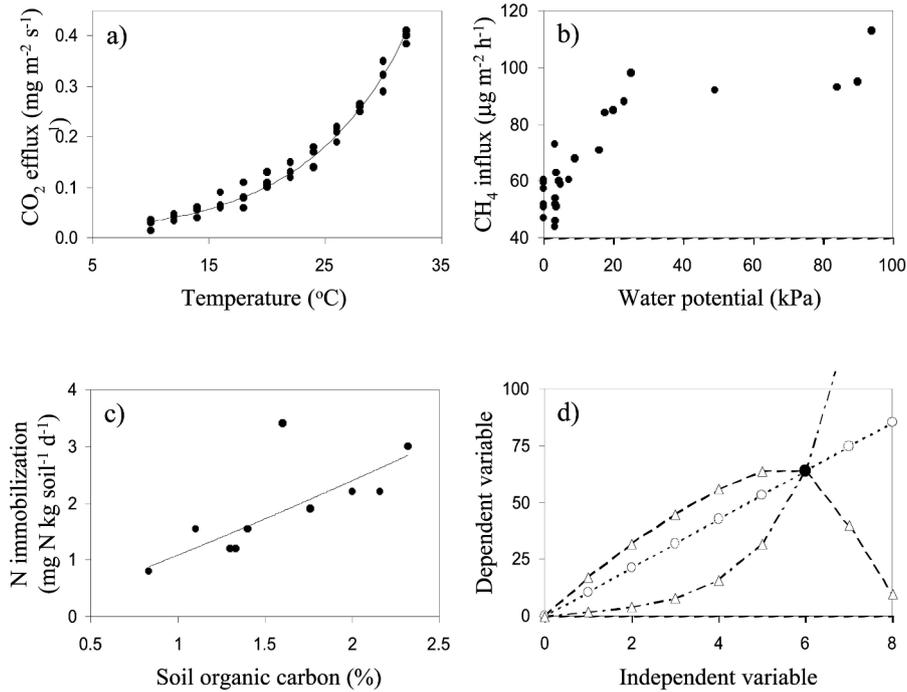

Figure 6.2. *Examples of nonlinear relationships between a process (Y) and regulatory factor (X). For the purposes of our discussion, nonlinearity exists when the relationship between X and Y is not constant (i.e., dY/dX ≠ constant). (a) A positive, exponential relationship exists between respiration from a forest soil and soil temperature across 10 to 32°C. Redrawn from Fang and Moncrieff (2001); not all single measurements (filled circles) are shown. (b) Bradford et al. (2001) found that soil water potential explains 78% of the variation in net, soil methane oxidation rates measured across one year in a temperate forest. (c) Barrett and Burke (2000) report a strong linear correlation between soil organic carbon content and gross nitrogen immobilization but in fact the relationship is better described by a power function (shown as solid line). (d) Examples of relationships (broken lines with open symbols) that could be constructed were measurements (filled circle denotes the mean) of a process (y-variable) taken at only one mean value of the regulatory factor (x-variable), such as if plots were sized to capture the spatial heterogeneity in regulatory factors within single plots (see Figure 6.1a). With only a single x-value there is no way to establish the shape of the relationship between X and Y.*

Theoretically, an alternative approach – to establish plots within spatially homogenous patches (Figure 6.1b) – generates the same rate estimates as obtained using larger plots and enables the shape of the relationship between process and regulatory factor to be quantified. A second advantage is that smaller plot sizes are often more tractable for manipulation but the trade-off is that greater replication is needed. Selection of a plot size that falls somewhere between (Figure 6.1c), the size that captures the spatial heterogeneity in a single plot (Figure 6.1a) and smaller plots that contain homogenous areas (Figure 6.1b), can only inform about the response of the specific experimental plot being investigated; the experimental results will have little quantitative relevance to areas outside of the experimental plot.

Our discussion of what constitutes a suitable plot size for valid spatial extrapolation belongs in an "ideal world" – in reality, a number of complications come to bear. First, large plots that capture spatial heterogeneity will rarely have smaller homogenous areas evenly proportioned within them and, therefore, rate estimates may differ between large and small plot experiments. Appropriate replication can reduce these differences. Second, single factors do not regulate biogeochemical processes although, in many cases, a single factor can explain a significant part of the observed variation in measured process rates (e.g., Figure 6.2a-c); in such instances, basing plot size on the spatial heterogeneity of this factor will provide more information than if spatial considerations were ignored. Third, spatial patterns may vary with time. This limitation is perhaps most relevant when variation is stochastic (e.g., localized rainfall within a desert) as opposed to non-stochastic (nutrients accumulate under desert shrubs across time, increasing differences in soil fertility between shrubs and interspaces; Schlesinger et al. 1996). Fourth, the identity of the landscape unit next to the one of interest, and/or the degree of connectedness between similar units (e.g., corridors in fragmented landscapes), can markedly influence processes (see Peters et al., Chapter 7, Groffman et al., Chapter 10, Lloyd et al., Chapter 14). Fifth, in many experiments multiple processes are measured within an experimental plot (e.g., Shaver et al. 1998, DeLucia et al. 1999, Bradford et al. 2002) and their regulatory factors may not only be different but also distributed, in terms of their heterogeneity, across distinctly different spatial areas. As a result there may not be a single plot size that enables all the measured processes to be extrapolated legitimately. And finally, in many cases there may be no *a priori* knowledge of which factor(s) regulates a process until an experiment has been conducted.

The above scenarios illustrate the need to consider the appropriate plot size in designing IEE studies. For example, one may decide that issues of spatial scale cannot be incorporated into an experiment due to limiting resources or knowledge about regulatory factors. If so, at least this is a conscious decision and as such the ability to extrapolate the results spatially can be critically assessed. Such "scale awareness" is a marked improvement on the current norm (Gardner et al. 2001).

6.3.2 Variability in External Regulatory Factors

In this section we develop our discussion of spatial scale in IEEs. What temporal scale considerations are necessary when biogeochemical processes are primarily

regulated by external factors, and secondarily by internal factors? To illustrate, we use an excellent IEE example in the recent literature where both *temporal* and *spatial* patterns are shown to determine the absolute and relative magnitude of treatment effects. Smith et al. (2000) exposed an intact Mojave Desert ecosystem to elevated atmospheric CO_2 and showed that new shoot production of a dominant perennial shrub doubled in a high rainfall year but was not affected by elevated CO_2 in the subsequent year, which was characterized by low rainfall. Further, in the high rainfall year production of annual and exotic grasses, which failed to germinate in the drought year, was higher under elevated CO_2 and the magnitude of the treatment response was greater within fertile "shrub-islands" than in open interspaces between perennial plants. Precipitation – an external factor that varies temporally – is the dominant control on system productivity and when delivered in sufficient amounts the spatial pattern of soil fertility across the landscape (an internal factor) becomes an important determinant of the system response to elevated CO_2. Such information on the spatial and temporal controls on productivity is necessary to predict how the Mojave might respond, at the landscape level, to future environmental change.

Spatial and temporal patterns in deserts are of course typified by extremes. Identifying and then incorporating similar scale considerations into other types of ecosystems (e.g., tropical and temperate forests) that appear more spatially and temporally homogenous in terms of factors, such as soil nutrient availability and rainfall, may be more difficult but potentially no less pertinent. For example, the magnitude of plant biomass responses to elevated CO_2 appears dependent on soil nutrient availability for the majority of species tested from across biomes (Bazzaz and Catovsky 2002). The primacy of this control is likely to be dependent on other factors, such as climate (e.g., Smith et al. 2000), but commonly there is detailed information on variation in such factors across years, which can be used to determine the degree of confidence that one can extrapolate results in time and space. However, published relationships between processes and regulatory factors should be assessed with caution prior to designing spatially- and temporally-sound experiments. For example, actual evapotranspiration (AET) explains marked variation in plant litter decomposition rates across latitudes (Aerts 1997), but within a particular system litter quality or soil community composition may be a better predictor (Aerts 1997, González and Seastedt 2001). Were such scale considerations to be ignored, predictions about how decomposition will respond to perturbation might very well be flawed. Probably the best example of such scale-ignorance is the plant diversity-productivity debate. Here, the hump-shaped relationship of diversity with increasing productivity (see Mittelbach et al. 2001), generated from inter-community/site comparisons across a landscape, has been used to suggest that a similar relationship should occur within a single community/site (Loreau et al. 2001).

6.3.3 Temporal Scale

Many of the same considerations used to determine appropriate spatial scales for extrapolation are relevant when deciding upon the necessary temporal grain and extent to extrapolate experimental findings. After defining *grain* and *extent*, we

discuss these considerations with respect to temporal scale and at the same time continue discussion appropriate to spatial scaling.
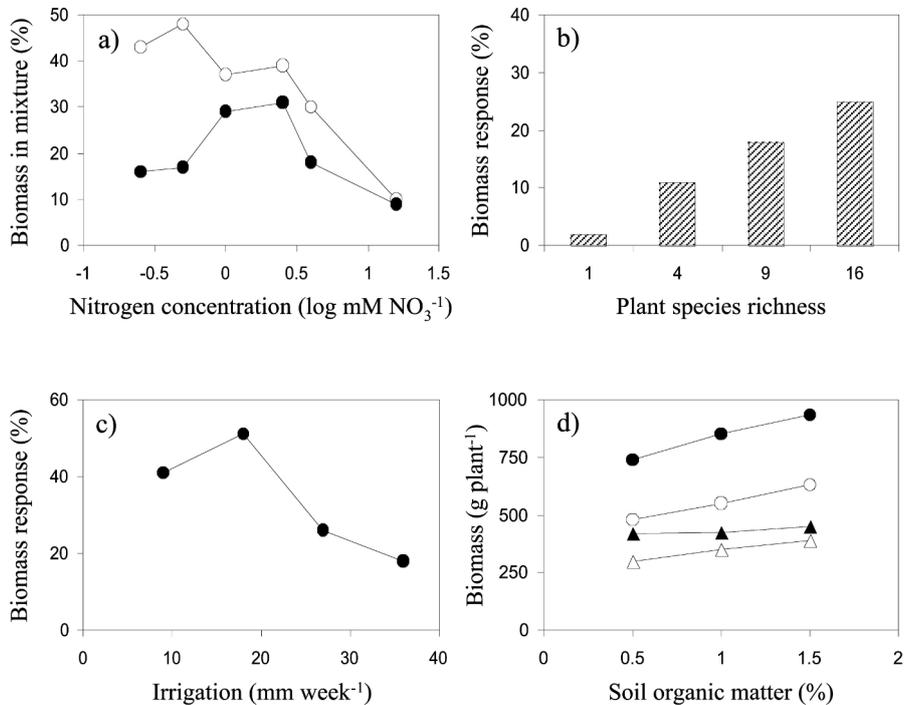

Figure 6.3. *Examples of nonlinear relationships between the magnitude of treatment effect on a process and different levels of the regulatory factor. (a) Effect of CO_2 concentration (ambient, open circles; elevated, filled circles) on the proportion of biomass that is Trifolium repens in four species plant mixtures, across a gradient of nitrogen availability. At low nitrogen concentrations there is a marked effect of CO_2 but at high nitrogen, the CO_2 concentration does not affect representation of T. repens in the mixtures (Navas et al. 1999). (b) The impact of nitrogen addition (compared to the control) on the biomass response of plant communities of increasing species richness. More diverse communities show a greater response to nitrogen treatment (Reich et al. 2001). (c) Volk et al. (2000) show that the relative response of plant biomass in two species mixtures to an increase in CO_2 concentration markedly decreases when amended with higher levels of simulated rainfall. (d) Soil organic matter content (a surrogate for soil fertility) regulates the aboveground biomass of individual, two-year old, loblolly pine seedlings but planting density moderates its regulatory role (Zutter et al. 1997). One (filled circles), two (open circles), three (filled triangles) and four (open triangles) seedlings were planted per m^2.*

Grain is the smallest spatial or temporal resolution at which data are collected (i.e., the limit to which individual measurements can be resolved), whereas extent is the spatial size (length, area or volume) of the study and the total duration over which it is observed (Wu and Li, Chapter 1). See Kemp et al. (2001) for discussion

of how these definitions are modified depending on whether the data of interest are observed in nature, collected through experimental manipulations, or measured as intrinsic scales of a natural system. In the previous section we used the Smith et al. (2000) study to highlight how temporal variability in regulatory factors can often affect the magnitude of treatment response of a process. This builds on the argument that relationships between processes and regulatory factors may be nonlinear (Figure 6.2) because it can be demonstrated that a linear relationship between the magnitude of treatment effect on a process and the regulatory factor(s) is unlikely (Figure 6.3). Thus, the temporal extent of IEE studies must capture enough variability to construct a robust relationship between regulatory factors, processes and treatment effects. This will be dependent on (i) variability in the regulatory factor(s), (ii) the process measured, (iii) the presence of synergies, feedbacks and thresholds, and (iv) how each treatment is applied. We discuss each in turn, highlighting the four main areas that we believe are necessary for consideration when deciding the temporal grain and extent of an experiment.

6.3.3.1 Variability in regulatory factors

Similar to selection of plot size (i.e., spatial extent), there is a strong relationship between the amount of heterogeneity in the regulatory factor that the experimenter is able to capture (to explain variation in the measured process) and the robustness of the mechanistic understanding generated. That is, the more levels of the regulatory factor at which the process is measured, the better quantified the relationship. Ideally, experiments should be run for as long as is feasible to enable this. When choosing spatial extent one can measure the spatial arrangement of regulatory factors; however, in setting the temporal extent one must presuppose how the regulatory factor will vary over time. This might be based on past records (if they exist), speculations, hunches, analogies with other studies, etc.: in other words, predict the future! Here, perhaps, is when the experimenter is most at the mercy of factors outside experimental control.

6.3.3.2 Frequency of measurement: integrative vs. active processes

The frequency with which to measure a process is the next decision to be made. The term *frequency* is synonymous with that of *sampling density*. The frequency chosen for a study will be dependent upon the process of interest (see below) and there is little to be gained by setting the frequency to the same resolution as the temporal extent of the IEE. For example, if we measured tree growth once at the end of a three year experiment it is unlikely that we could ever extrapolate this result in time. We would need exactly the same pattern of climatic conditions to manifest over a three year period to do so, which is unlikely. However, had we increased the frequency of our measures we would have been able to plot the treatment effect against values for regulatory temporal factors (e.g., climatic variables) and therefore construct a relationship potentially useful for extrapolation – as done by Smith et al. (2000). These authors measured aboveground production of the dominant shrubs at

their desert site in two successive years that differed markedly in rainfall and, as a consequence, rate of shrub growth. Had they measured production only at the end of the second year, then they could not have determined that production was dependent upon rainfall within a year because they would have had a single production estimate that integrated across two years of different rainfall patterns.

Commonly, the process of interest determines the frequency that we choose. For example, estimates of net primary productivity (NPP) constitute a measure of plant biomass production across time that is influenced by external factors such as climate and nitrogen deposition, and by internal factors such as site fertility, herbivory, competition, species identity and facilitation. All of these factors operate over the growing season but NPP is usually measured only once (e.g., during peak growth) or twice (rarely) to assess treatment effects (e.g., Hector et al. 1999, Shaw et al. 2002). We define measures such as NPP as a *temporally-integrative* process, in contrast to a *temporally-active* process, such as soil CO_2 efflux. The former may have units of, for example, $kg\ m^{-2}\ y^{-1}$ whereas the latter, $\mu mole\ m^{-2}\ s^{-1}$. These distinctions muddy the notion of frequency to a certain extent because it does not simply define how often a process is measured. For example, if we measure NPP and soil CO_2 efflux once a year then the frequency is the same; however, the information gathered regarding the response of the process to the treatment will reflect very different time-scales (and therefore temporal grains!). Further, the integrative vs. active nature of a process is likely to be on a continuous rather than discrete scale, and the exact positioning for a process may depend on what ecosystem is being investigated. For example, soil CO_2 efflux in a well-drained soil may be a truly active process, but in a wetland, where movement of gas through soil is restricted, changes in atmospheric pressure and freeze-thaw cycles can lead to release of large build-ups of trace gases (Bubier et al. 2002). Hence the active vs. integrative nature of a specific process may be temporally, as well as ecosystem, dependent.

Such considerations directly impinge on the frequency with which we decide to measure a process and what mechanistic understanding we construct. For temporally-active processes, which have the potential to exhibit marked short-term variation if regulated by, say, a single but temporally dynamic factor, frequent quantification is necessary to generate robust estimates (e.g., Bradford et al. 2001). The advantage is that frequent quantification will permit strong relationships between process and regulator to be constructed within relatively short time-scales. In contrast, the mechanistic basis for variation in a temporally-integrative process such as NPP can be difficult to obtain because so many factors have had the opportunity to affect the rate estimate obtained at a single time point. Even after a number of years of measurement, causative mechanistic-relationships may be hard to identify (Shaw et al. 2002, Field et al. 2003, Mitchell and Reich 2003), unless regulatory factors exert their influence in a temporally predictive manner (e.g., litter decomposition shifting from nitrogen- to lignin-controlled as in Taylor et al. 1989). Identification of causation may be further compounded by changes induced by treatment on the system of study over time that then interact with the treatment to modify the original treatment impacts on the process of interest, as discussed next.

6.3.3.3 Synergies, feedbacks and thresholds

The magnitude and direction of treatment effects in IEEs may change over time not simply through variation in the magnitude of a regulatory factor (Figure 6.3) but because of synergies, feedbacks and thresholds (Reynolds et al. in press). Synergies arise when two or more processes/factors interact in such a way that the outcome is greater than the sum of their separate effects (e.g., multiplicative rather than additive). Feedbacks occur when a system or process is modified by changes in its own influence or size. If the initial (direct) response is enhanced, the feedback is considered positive; if decreased, the feedback is negative. Thresholds occur where critical values or set points are reached. They often result in major nonlinear changes in a process, and are products of the complex of synergies and feedbacks that exist in both climate and biological systems (Harte 1996). Thus, in short, the temporal responses of processes to treatment will be the result of variation in regulatory factors modified by the action of synergies, feedbacks and thresholds (examples are provided in Table 6.1) – a highly complex situation indeed!

How one can reconcile the potential complexity of mechanisms that result in observed treatment effects with the necessary understanding for temporally predictive science is a moot point. Schellnhuber (1999) challenges us to consider: "Where are the limits of scientific predictability in complex systems?" Holling et al. (2002) argue that it is essential to distinguish what is understandable or predictable (even if uncertain) from that which is inherently unpredictable. It is clear that prediction (i.e., temporal extrapolation) operates at the frontier of what is amenable to the scientific method today and this must be borne in mind when questions of temporal scaling are posed.

6.3.3.4 Press vs. pulse and step vs. gradual

Intact ecosystem experimental treatments can be categorized as press or pulse. Based on the definition of Bender et al. (1984), a *press* experiment is where the perturbation is sustained (often at a constant level) across time (e.g., exposure of intact communities to elevated CO_2). In contrast, *pulse* experiments involve a relatively brief period of perturbation (e.g., single application of nitrogen fertilizer). The investigator is typically interested in how a system/process deviates from its pre-treatment state (press and pulse) and then how quickly, if at all, it returns to its pre-treatment state (pulse only). With pulse experiments the transient response of a process to treatment is usually the response of most interest. For accurate quantification of process values across time, knowledge of the timing of process response to treatment is required. Williams et al. (1999) demonstrated that N_2O efflux from grassland soils amended with bovine urine was practically instantaneous (<4 h) and that the flux values observed were approximately six times higher within the first 24 hours than after them. This marked initial response was missed in earlier studies, leading to underestimates of soil N_2O flux. Where more integrative processes are measured (e.g., Pfisterer and Schmid 2002), a single process measurement may integrate across declining or increasing responses of the process to treatment over time but at a temporal grain finer than that measured. Thus, the

sensitivity of the measured process response to the pulse perturbation may be under- or over-estimated, respectively.

Table 6.1. *Examples of ecological nonlinearities on biogeochemical processes caused by thresholds, feedbacks and synergisms.*

Thresholds	*Feedbacks*	*Synergisms*
Al stress-acid deposition: The calcium/aluminum (Ca/Al) molar ratio of the soil solution is an ecological indicator of thresholds, beyond which the risk of forest damage from Al stress and nutrient imbalances increases (Cronan and Grigal 1995). Based on data from 89 forested catchments across Europe, Dise et al. (2001) found thresholds for N and S depositions, above which Al is released from forests at maximum rates.	*Herbivores-plants-N cycling*: Herbivores indirectly decelerate N cycling in savannah by decreasing the abundance of plant species with nitrogen-rich tissues. They may also decelerate succession by reducing the cover and biomass of woody plants (Ritchie et al. 1998).	*Vegetation pattern-NPP*: In arid ecosystems clumping or banding of vegetation results in an overall increase of production and plant diversity due to the combined benefits of the redistribution of precipitation and nutrients into patches (Aguiar and Sala 1999).
Weed density-productivity: When weed populations reach a threshold density, crop production is reduced (Brown et al. 1999).	*Elevated CO_2-respiration*: Plants grown in elevated CO_2 increase C input to the rhizosphere, which in turn feedbacks to affect tree growth and soil C dynamics. For example, additional soil C inputs increase rhizosphere respiration (Cheng 1999).	*Litter quality-decomposition*: Mixing litters of different nitrogen concentration has negative synergistic effects on decomposition rate (Smith and Bradford 2003).
Canopy chemistry-N cycling: Estimates of soil C:N ratios indicate that 63% of the White Mountains in New Hampshire fall below 22, a critical threshold for the onset of nitrification (Ollinger et al. 2002).	*Canopy chemistry-N cycling*: Foliar N increases with increasing net N mineralization, suggesting a positive feedback between foliar chemistry and soil N status (Ollinger et al. 2002).	*Insects-fire-C sequestration*: Climatic change will affect the frequency and intensity of pest outbreaks, which causes a considerable loss of wood and affects fire frequencies. This synergism between outbreaks, fires, and warming must be accounted for in models if realistic carbon
	Herbivore-root growth: Herbivores have a	

Table 6.1 (contd.)

Thresholds	Feedbacks	Synergisms
Plant cover-soil loss: Minor changes in shrub or grass cover resulting from climatic variability and grazing can lead to sudden increases in soil erosion (Breshears and Barnes 1999). *Bark thickness-fire susceptibility*: Forest fragmentation and intentional burning of grasslands in the dry, subhumid tropical forests of Bolivia is increasing. Hence, the frequency of fires in local forests, which did not develop with fires, is likely to increase. A bark thickness below 18 mm makes the tree species highly susceptible to fire damage and 81% of the species are below this threshold and, therefore, likely to experience cambial injury from low intensity fires; thus, biomass C storage in this region is at grave risk (Pinard and Huffman 1997).	positive feedback on root growth, stimulating growth 7× more than shoot production (Frank et al. 2002). *Logging-fires-C loss*: For the 1997-98 Indonesian fires, the largest fire disaster ever observed, forest fires primarily affected recently logged forests while primary forests or those logged long ago were less affected. These results support the hypothesis of positive feedback between logging and fire occurrence (Siegert et al. 2001), both of which affect C loss from tropical forests. *Soil N-plant growth-CO_2*: Populus grown under elevated CO_2 and low-N soil have reduced photosynthetic capacity, but in high-N soil, photosynthesis is enhanced (Zak et al. 1993).	sequestration forecasts in a warming climate are to be made (Volney and Fleming 2000). *Plant diversity-soil processes*: Two-species mixtures of ten different grassland plants had higher or lower soil microbial biomass, respiration: biomass ratio and plant litter decomposition than the additive prediction from monoculture treatments (Wardle and Nicholson 1996). *Soil nutrients-UV radiation-plant growth*: Enhanced UV-B radiation (equivalent to a 15% ozone depletion) and high soil nutrient levels interact synergistically to accelerate growth of *Phlomis fruticosa*, a fast-growing, drought-deciduous Mediterranean plant species (Levizou and Manetas 2001).

Given that the main focus of pulse experiments is to quantify transient dynamics, they tend to be less criticized than press experiments for being short-term. However, the criticism that short-term experiments may quantify transient dynamics which may be opposite to long-term effects (Tilman 1989), is equally applicable to press and pulse manipulations. These "transient dynamics" may result from ecological mechanisms that would be observed under natural circumstances, such as the crossing of critical thresholds (see Table 6.1 for examples). Alternatively, they may be artifacts of the way in which treatment is applied. For example, in the Duke

Forest FACE study, where an entire stand of loblolly pine is being exposed to elevated atmospheric CO_2 concentrations, the growth of the forest in the first three years was significantly enhanced. However, after three years this stimulation was lost, a result of the negative feedback of insufficient soil nutrients (Oren et al. 2001). In a modeling exercise, parameterized using data from the same FACE site, Luo and Reynolds (1999) demonstrated that had the treatments been applied in a gradual increment in CO_2 concentration – rather than a step from ambient to 550 ml l^{-1} – then the marked and transient increase in productivity would not have been observed. This was due to differences in forest N demand between a gradual and step increase in CO_2. Of course, many treatments are applied in such a "step" fashion – we simply do not have the time to "realistically" mimic a gradual change in CO_2 concentration, mean annual temperature, or rainfall variability, as predicted under IPCC scenarios (Houghton et al. 2001). What is important is to recognize the implications of such step changes on our ability to extrapolate results in time and how the mechanistic understanding generated aids future prediction (e.g., Luo 2001). For some treatments the step vs. gradual categorizations are less relevant, if at all. For example, one can realistically impose treatments to investigate the response of an agricultural field to long-term fertilizer application (Hütsch et al. 1993) or the novel occurrence of exotic species within ecosystems (Alvarez and Cushman 2002). Regardless of whether the "step" categorization applies, the whole concept of step vs. gradual application has stimulated important debate and experimentation into the effects of different levels of applied treatment.

The responses of processes to different levels of treatment are typically nonlinear, even when applied over the same spatial and temporal extent. For example, in a Texas grassland net N mineralization rates in soils decrease exponentially with increasing CO_2, with the largest changes occurring at subambient concentrations (Gill et al. 2002). In a separate study, Granados and Körner (2002) examined the potential for elevated CO_2 to accelerate the growth of climbing vine species of tropical forests, and hence tree turnover. In both species, the increase from 280 (pre-industrial) to 420 ml l^{-1} CO_2 had a much larger effect on growth than the increase from 420 to 560 ml l^{-1}. Further, the 560 to 700 ml l^{-1} enhancement caused a massive reduction of the stimulation. Yet most experiments use only one treatment level, potentially masking nonlinear relationships. If we wish to extrapolate treatment effects in time or space, across which the magnitude of the press or pulse of the perturbation simulated by the experimental treatment may differ, then we need to quantify the shape of the relationship between processes and treatment levels.

6.4 MEASURING TREATMENT EFFECTS USING MODEL ECOSYSTEM EXPERIMENTS

6.4.1 Laboratory and Field MEEs

We define MEEs as those where the experimental system is either wholly or partially constructed. Our definition includes three types of MEEs. In type (1) the

climate is constructed and the biotic component remains "intact" – i.e., laboratory experiments using, for example, vegetated soil monoliths that are removed from the field. In type (2) the climate is "intact" and the biotic component is constructed – i.e., field experiments where, for example, the plant community is grown from seed and controlled for richness and functional-type composition. Lastly, in type (3) both the climate and the biotic component are constructed. These are laboratory experiments using biotic systems that are not "intact", e.g., those typically conducted in controlled environment chambers (e.g., Lawton 1995, Reynolds 2001).

Type (1) MEEs isolate an "ecosystem-unit" from the rest of its landscape and introduce a suite of questions regarding the "reality" of process rates and treatment effects measured under laboratory conditions. For example, what impact will the absence of immigration have? Has the disturbance altered soil nutrient availabilities? Do diurnal patterns of temperature need to be simulated? Questions with respect to disturbance effects of the removal into the laboratory have a different basis to those that question the effects of obstructing processes such as immigration and diurnal variation in climate. Disturbance effects are non-target changes in the system of interest, whereas isolation from biotic and abiotic factors that would not be under experimental control in intact ecosystems is a deliberate attempt to maximize the relative contribution of treatment to variation in measured process rates. In short, the aim is to increase the chances of detecting a signal, if one exists, by reducing noise.

The rationale of maximizing the signal-to-noise ratio also underlies the use of type (3) MEEs. In addition, many of the treatments imposed on these MEEs may not be feasible under field conditions. For example, to manipulate the soil community in the field, and maintain it over biological meaningful temporal and spatial scales, presents many logistic difficulties (Kampichler et al. 2001). Thus, partial or complete construction of an experimental system is often necessary to test the posed hypotheses (e.g., the model soil communities created by Bradford et al. 2002 and De Deyn et al. 2003). The same "feasibility" criterion underlies the use of the majority of field MEEs. For example, construction of plant communities of set initial density and varying diversity simply would not be possible without a constructionist approach (e.g., Hooper and Vitousek 1997). Given that the biotic system has been partially constructed, which clearly is not a natural situation, the decision to conduct the experiment in the field rather than in the laboratory is often an attempt to increase the realism of the experimental system. These constructed vs. real components to MEE design are properties associated with internal and external validity.

6.4.2 Internal and External Validity

Manly (1992) stated that, "Internal validity concerns whether the apparent effects or lack of effects shown by the experimental results are due to the factor being studied, rather than some alternative factor. External validity concerns the extent to which the results of an experiment can be generalized to some wider population of interest." Internal validity is likely to be highest when all factors are under the experimenter's control. External validity will be highest when the experimental system corresponds to the natural system under investigation in all aspects (Naeem

2001) and, following Manly's definition, when the experimental plot is appropriately temporally and spatially scaled for extrapolation. Given the huge number of studies conducted using model ecosystems, and the fact that they all trade-off external for internal validity to some extent, it is appropriate to ask: What understanding we can extrapolate from MEEs to facilitate temporally and/or spatially predictive science? Maybe this very question is why some of the most well known MEEs, such as those that initially evaluated the effects of elevated CO_2 on ecosystem response (e.g., Billings et al. 1982) and the role of biodiversity in ecosystem functioning (e.g., Naeem et al. 1994), have produced much debate and experimentation to challenge/confirm the original findings.

6.4.3 Scaling MEEs – Process Rates and Treatment Effects

Our discussion of the factors that must be considered prior to extrapolating experimental results in space or time from intact systems apply equally to MEEs (see Gardner et al. 2001 for new perspectives and review of this topic). Given these obstacles, and the question of validity, when can we extrapolate process rates and absolute/relative treatment effects from MEEs? The answer is probably "almost never" (if we decide that we can, then uncertainty analysis will be an essential component; see Li and Wu, Chapter 3).

To support our contention that process rates and absolute/relative treatment effects from MEEs can almost never be extrapolated we provide the following hypothetical example. A type (1) MEE (intact system, constructed climate) is used by Dr. Climatron to investigate how soil CO_2 efflux is affected by carbon addition. Prior fieldwork by his group has demonstrated that soil moisture and temperature are regulators of CO_2 efflux and so they collect a soil monolith suitably sized to capture heterogeneity in moisture at the landscape level. Glucose solution amendments are imposed as the treatment and water amendments as the control. Carbon dioxide efflux is then measured across a range of soil moistures at different temperatures. Using field measurements of temperature, moisture and efflux, from carbon amendment and control plots, Dr. Climatron's group cross-calibrate the relationship generated in the laboratory with that in the field. They find that temperature and moisture explain most of the field variation in efflux and, as a result, can quite accurately predict the measured CO_2 efflux of the monoliths maintained in the laboratory from the imposed soil moistures and temperatures. That is, their laboratory measurements have high external validity. They then use their MEE to determine the efflux at temperatures and moistures higher and lower than those observed during their field observations, and at different levels of carbon amendment. Thus, they are quantifying relationships at values beyond those calibrated against field measurements.

This property of MEEs is a strength of the MEE approach, but how confident can we be in their extrapolation? Given the high external validity of their original laboratory measurements (given the good fit with field data) our confidence is high but we recognize that they are assessing *potential* and not *actual* process rates. Hence, whether we can scale their results to the field becomes questionable. Had they first sieved the soil and then reconstructed the monolith, before placing it in the

field (type [2] MEE) or laboratory (type [3] MEE), they would still be measuring potential effects but our confidence in whether the measured process rates reflect those of an intact system would be much less than when they used a type (1) MEE. For example, they may have altered characteristics of the system, such as drainage, through sieving which might influence efflux. Nevertheless, carbon amendment would probably still stimulate heterotrophic respiration, so we might be confident to extrapolate the sign of the treatment effect, which would be positive, to the field.

We suggest that process rates and absolute/relative treatment effects measured in MEEs should not be extrapolated to intact systems. Further, through controlling for regulators and isolating the system in the case of lab MEEs, we remove mechanisms that may contribute to feedbacks, synergisms and thresholds. Thus, the longer we run a study, or the simpler a system is, then the more we must question whether the sign of the treatment effect can be extrapolated. It is worth noting that the sign of treatment effects in relatively long-term and complex MEEs are the same when both field and laboratory MEEs are used (Naeem et al. 1994, Hector et al. 1999), despite marked differences in validity. In contrast, using simpler systems, Navas et al. (1999) demonstrated that type (3) MEEs with isolated plants could not be used to predict the behavior of species mixtures under elevated CO_2 and a nitrogen gradient, whereas monocultures could. Their work suggests that there may be critical levels of validity across which we cannot scale. It is imperative that we determine if such critical thresholds of validity exist given the very large number of MEEs conducted using single plants.

Naeem (2001) provides a more in-depth evaluation of how validity relates to extrapolation and, interestingly, concludes that the current biodiversity-productivity debate is premature given that appropriately valid/scaled experiments have not been conducted to address the question. His conclusion may be applicable to many scientific debates because of the lack of "scale awareness" in the experiments conducted by all sides to support their arguments.

6.4.4 Scaling MEEs – Mechanisms

So when are MEEs useful? We believe MEEs are best suited to identifying causation/mechanisms and become most powerful in this respect as internal validity is increased. For example, demonstration of semi-conservative replication of DNA would not have been possible without a highly simplified MEE (Meselson and Stahl 1958). Similarly, Goddard and Bradford (2003) used the MEE approach to demonstrate the potential for a population of fungi found widely in terrestrial ecosystems to evolutionarily adapt, within relatively few generations, to altered carbon and nitrogen availability in the environment. Further, the role of soil fertility in determining plant responses to elevated CO_2 was worked out using MEEs (Bazzaz and Catovsky 2002). However, at the same time internal validity increases such that our confidence in extrapolating this causation to natural systems decreases due to loss of external validity. In a thought provoking article Lawton (1995) challenges us to consider this loss as a research question as opposed to a limitation. We agree. For the purposes of extrapolation of mechanisms from MEEs to the field, we suggest that short-term, complex MEEs are best suited. If we wish to know how

important a mechanism may be in an intact system then, given that as complexity increases the signal-to-noise ratio decreases, one approach may be to identify a potential mechanism using a highly simplified system. Then, while maintaining the same temporal extent, increase system complexity and measure whether the mechanism of causation is still detectable. Unfortunately there is still a catch – weak interactions, which are hard to detect above non-controlled variation, are increasingly being shown to be important in ecological dynamics (e.g., Berlow 1999).

6.5 CONCLUSIONS

The challenges of global changes and Earth System Science are great. To extrapolate biogeochemical process rates, and treatment effects on them, across time and space greater than the experimental extent requires an understanding of the factors that regulate the process, how these factors are temporally and spatially distributed, and what feedbacks, thresholds and synergies may manifest. We cannot expect the shape of the relationship between process rates, or treatment effects, and regulators to be linear; multiple levels of treatments must be imposed across varying values of regulatory factors. We need to determine if the process being measured is integrative or active within the specific ecosystem under investigation and whether the treatment responses observed are artifacts of the way treatment is imposed (e.g., step vs. gradual).

Spatial and temporal extrapolation differs in a key regard. The latter attempts to predict the future behavior of ecosystems and is inherently more complex due to the greater number of uncertainties and interactions that act across temporal as opposed to spatial scales. Research intent on extrapolation might then be best focused on spatial issues but, given the societal need for temporally predictive science, well-resourced multi-disciplinary studies that critically evaluate their own limitations will be required to tackle temporal scaling. Model ecosystem experiments will contribute to this effort and their use will be best targeted at unravelling the mechanisms behind causation of treatment effects observed in intact systems. Neither intact systems nor model systems alone will provide the necessary understanding required to scale experimental results across time and space, and discussion (Carpenter 1996, Carpenter 1999, Drenner and Mazumder 1999) of which is superior is constructive only if to highlight this point. There is a need for much greater "scale awareness" in ecology and this is reflected both in this and related volumes (e.g., Gardner et al. 2001).

ACKNOWLEDGEMENTS

REFERENCES

Aerts, R. 1997. Climate, leaf litter chemistry and leaf litter decomposition in terrestrial ecosystems: a triangular relationship. Oikos 79:439-449.

Aguiar, M. R., and O. E. Sala. 1999. Patch structure, dynamics and implications for the functioning of arid ecosystems. Trends in Ecology and Evolution 14:273-277.

Alvarez, M. E., and J. H. Cushman. 2002. Community-level consequences of a plant invasion: effects on three habitats in coastal California. Ecological Applications 12:1434-1444.

Barrett, J. E., and I. C. Burke. 2000. Potential nitrogen immobilization in grassland soils across a soil organic matter gradient. Soil Biology & Biochemistry 32:1707-1716.

Bazzaz, F. A., and S. Catovsky. 2002. Plants – from cells to ecosystems: impacts of global environmental change. Pages 94-111 *in* H. A. Mooney and J. G. Canadell, editors. Encyclopaedia of Global Environmental Change. Wiley, Chichester, UK.

Bender, E. A., T. J. Case, and M. E. Gilpin. 1984. Perturbation experiments in community ecology: theory and practice. Ecology 65:1-13.

Berlow, E. L. 1999. Strong effects of weak interactions in ecological communities. Nature 398:330-334.

Billings, W. D., J. O. Luken, D. A. Mortensen, and K. M. Peterson. 1982. Arctic tundra: a source or sink for atmospheric carbon dioxide in a changing environment. Oecologia 53:7-11.

Bradford, M. A., T. H. Jones, R. D. Bardgett, H. I. J. Black, B. Boag, M. Bonkowski, R. Cook, T. Eggers, A. Gange, S. J. Grayston, E. Kandeler, A. E. McCaig, J. N. Newington, J. I. Prosser, H. Setälä, P. L. Staddon, G. M. Tordoff, D. Tscherko, and J. H. Lawton. 2002. Impacts of soil faunal community composition on model grassland ecosystems. Science 298:615-618.

Bradford, M. A., P. A. Wookey, P. Ineson, and H. M. Lappin-Scott. 2001. Controlling factors and effects of chronic nitrogen and sulphur deposition on methane oxidation in a temperate forest soil. Soil Biology & Biochemistry 33:93-102.

Breshears, D. D., and F. J. Barnes. 1999. Interrelationships between plant functional types and soil moisture heterogeneity for semiarid landscapes within the grassland/forest continuum: a unified conceptual model. Landscape Ecology 14:465-478.

Brown, J. R., J. Herrick, and D. Price. 1999. Managing low-output agroecosystems sustainably: the importance of ecological thresholds. Canadian Journal of Forest Research 29:1112-1119.

Bubier, J., P. Crill, and A. Mosedale. 2002. Net ecosystem CO_2 exchange measured by autochambers during the snow-covered season at a temperate peatland. Hydrological Processes 16:3667-3682.

Carpenter, S. R. 1996. Microcosm experiments have limited relevance for community and ecosystem ecology. Ecology 77:677-680.

Carpenter, S. R. 1999. Microcosm experiments have limited relevance for community and ecosystem ecology: reply. Ecology 80:1085-1088.

Cheng, W. X. 1999. Rhizosphere feedbacks in elevated CO_2. Tree Physiology 19:313-320.

Clark, J. S., S. Carpenter, M. Barber, S. Collins, A. Dobson, J. A. Foley, D. M. Lodge, M. Pascual, R. Pielke, Jr., W. Pizer, C. Pringle, W. V. Reid, K. A. Rose, O. Sala, W. H. Schlesinger, D. H. Wall, and D. Wear. 2001. Ecological forecasts: an emerging imperative. Science 293:657-660.

Committee on Global Change. 1988. Toward an understanding of global change. National Academy Press, Washington, D.C.

Cronan, C. S., and D. F. Grigal. 1995. Use of calcium/aluminum ratios as indicators of stress in forest ecosystems. Journal of Environmental Quality 24:209-226.

De Deyn, G. B., C. E. Raaijmakers, H. R. Zoomer, M. P. Berg, P. C. de Ruiter, H. A. Verhoef, T. M. Bezemer, and W. H. van der Putten. 2003. Soil invertebrate fauna enhances grassland succession and diversity. Nature 422:711-713.

DeLucia, E. H., J. G. Hamilton, S. L. Naidu, R. B. Thomas, J. A. Andrews, A. Finzi, M. Lavine, R. Matamala, J. E. Mohan, G. R. Hendrey, and W. H. Schlesinger. 1999. Net primary production of a forest ecosystem with experimental CO_2 enrichment. Science 284:1177-1179.

Dise, N. B., E. Matzner, M. Armbruster, and J. MacDonald. 2001. Aluminum output fluxes from forest ecosystems in Europe: a regional assessment. Journal of Environmental Quality 30:1747-1756.

Drenner, R. W., and A. Mazumder. 1999. Microcosm experiments have limited relevance for community and ecosystem ecology: comment. Ecology 80:1081-1085.

Fang, C., and J. B. Moncrieff. 2001. The dependence of soil CO_2 efflux on temperature. Soil Biology & Biochemistry 33:155-165.

Fernández, R. J., and J. F. Reynolds. 2000. Potential growth and drought tolerance of eight desert grasses: lack of a trade-off? Oecologia 123:90-98.

Field, C. B., M. R. Shaw, H. A. Mooney, E. S. Zavaleta, N. R. Chiariello, and E. E. Cleland. 2003. Assessing environmental changes in grasslands. Science 299:1844-1845.

Frank, D. A., M. M. Kuns, and D. R. Guido. 2002. Consumer control of grassland plant production. Ecology 83:602-606.

Gardner, R. H., W. M. Kemp, V. S. Kennedy, and J. E. Petersen, editors. 2001. Scaling Relations in Experimental Ecology. Columbia University Press, New York.

Gill, R. A., H. W. Polley, H. B. Johnson, L. J. Anderson, H. Maherali, and R. B. Jackson. 2002. Nonlinear grassland responses to past and future atmospheric CO_2. Nature 417:279-282.

Goddard, M. R., and M. A. Bradford. 2003. The adaptive response of a natural microbial population to carbon- and nitrogen-limitation. Ecology Letters 6:594-598.

González, G., and T. R. Seastedt. 2001. Soil fauna and plant litter decomposition in tropical and subalpine forests. Ecology 82:955-964.

Goudriaan, J., H. H. Shugart, H. Bugmann, W. Cramer, A. Bondeau, R. H. Gardner, L. A. Hunt, W. K. Lauenroth, J. J. Landsberg, S. Linder, I. R. Noble, W. J. Parton, L. F. Pitelka, M. Stafford Smith, R. W. Sutherst, C. Valentin, and F. I. Woodward. 1999. Use of models in global change studies. Pages 106-140 in B. H. Walker, W. L. Steffen, J. Canadell, and J. S. I. Ingram, editors. Global Change and the Terrestrial Biosphere: Implications for Natural and Managed Ecosystems. A Synthesis of GCTE and Related Research, IGBP Book Series No. 4. Cambridge University, Cambridge.

Granados, J., and C. Körner. 2002. In deep shade, elevated CO_2 increases the vigor of tropical climbing plants. Global Change Biology 8:1109-1117.

Harte, J. 1996. Feedbacks, thresholds and synergies in global change: population as a dynamic factor. Biodiversity and Conservation 5:1069-1083.

Hector, A., B. Schmid, C. Beierkuhnlein, M. C. Caldeira, M. Diemer, P. G. Dimitrakopoulos, J. A. Finn, F. H, P. S. Giller, J. Good, R. Harris, P. Högberg, K. Huss-Danell, J. Joshi, A. Jumpponen, C. Körner, P. W. Leadley, M. Loreau, A. Minns, C. P. H. Mulder, G. O'Donovan, S. J. Otway, J. S. Pereira, A. Prinz, D. J. Read, M. Scherer-Lorenzen, E.-D. Schulze, A.-S. D. Siamantziouras, E. M. Spehn, A. C. Terry, A. Y. Troumbis, F. I. Woodward, S. Yachi, and J. H. Lawton. 1999. Plant diversity and productivity experiments in European grasslands. Science 286:1123-1127.

Holling, C. S., L. H. Gunderson, and D. Ludwig. 2002. In quest of a theory of adaptive change. Pages 3-22 in L. H. Gunderson and C. S. Holling, editors. Panarchy: Understanding Transformations in Human and Natural Systems. Island Press, Washington D.C.

Hooper, D. U., and P. M. Vitousek. 1997. The effects of plant composition and diversity on ecosystem processes. Science 277:1302-1305.

Houghton, J. T., Y. Ding, D. J. Griggs, M. Noguer, P. J. van der Linden, X. Dai, K. Maskell, and C. A. Johnson, editors. 2001. Climate Change 2001: The Scientific Basis. Cambridge University Press, Cambridge.

Hütsch, B. W., C. P. Webster, and D. S. Powlson. 1993. Long-term effects of nitrogen-fertilization on methane oxidation in soil of the Broadbalk Wheat Experiment. Soil Biology & Biochemistry 25:1307-1315.

Kampichler, C., A. Bruckner, and E. Kandeler. 2001. The use of enclosed model ecosystems in soil ecology: a bias towards laboratory research. Soil Biology & Biochemistry 33:269-275.

Kemp, W. M., J. E. Petersen, and R. H. Gardner. 2001. Scale-dependence and the problem of extrapolation. Pages 3-57 in R. H. Gardner, W. M. Kemp, V. S. Kennedy, and J. E. Petersen, editors. Scaling Relations in Experimental Ecology. Columbia University Press, New York.

Lawton, J. H. 1995. Ecological experiments with model systems. Science 269:328-331.

Lawton, J. H. 1996. The Ecotron facility at Silwood Park: the value of "big bottle" experiments. Ecology 77:665-669.

Lawton, J. H. 2001. Earth system science. Science 292:1965.

Levizou, E., and Y. Manetas. 2001. Combined effects of enhanced UV-B radiation and additional nutrients on growth of two Mediterranean plant species. Plant Ecology 154:181-186.

Loreau, M., S. Naeem, P. Inchausti, J. Bengtsson, J. P. Grime, A. Hector, D. U. Hooper, M. A. Huston, D. Raffaelli, B. Schmid, D. Tilman, and D. A. Wardle. 2001. Biodiversity and ecosystem functioning: current knowledge and future challenges. Science 294:804-808.

Luo, Y. 2001. Transient ecosystem responses to free-air CO_2 enrichment (FACE): experimental evidence and methods of analysis. New Phytologist 152:1-8.

Luo, Y., and J. F. Reynolds. 1999. Validity of extrapolating field CO_2 experiments to predict carbon sequestration in natural ecosystems. Ecology 80:1568-1583.

Manly, B. F. J. 1992. The Design and Analysis of Research Studies. Cambridge University Press, Cambridge.

McGrady-Steed, J., P. M. Harrison, and P. J. Morin. 1997. Biodiversity regulates ecosystem predictability. Nature 390:162-165.

Meselson, M., and F. W. Stahl. 1958. The replication of DNA in *Escherichia coli*. Proceedings of the National Academy of Sciences of the United States of America 44:671-682.

Mitchell, C. E., and P. B. Reich. 2003. Assessing environmental changes in grasslands. Science 299:1844.

Mittelbach, G. G., C. F. Steiner, S. M. Scheiner, K. L. Gross, H. L. Reynolds, R. B. Waide, M. R. Willig, S. I. Dodson, and L. Gough. 2001. What is the observed relationship between species richness and productivity? Ecology 82:2381-2396.

Naeem, S. 2001. Experimental validity and ecological scale as criteria for evaluating research programs. Pages 223-250 *in* R. H. Gardner, W. M. Kemp, V. S. Kennedy, and J. E. Petersen, editors. Scaling Relations in Experimental Ecology. Columbia University Press, New York.

Naeem, S., L. J. Thompson, S. P. Lawler, J. H. Lawton, and R. M. Woodfin. 1994. Declining biodiversity can alter the performance of ecosystems. Nature 368:734-737.

Navas, M. L., E. Garnier, M. P. Austin, and R. M. Gifford. 1999. Effect of competition on the responses of grasses and legumes to elevated atmospheric CO_2 along a nitrogen gradient: differences between isolated plants, monocultures and multi-species mixtures. New Phytologist 143:323-331.

Neff, J. C., A. R. Townsend, G. Gleixner, S. J. Lehman, J. Turnbull, and W. D. Bowman. 2002. Variable effects of nitrogen additions on the stability and turnover of soil carbon. Nature 419:915-917.

Ollinger, S. V., M. L. Smith, M. E. Martin, R. A. Hallett, C. L. Goodale, and J. Aber. 2002. Regional variation in foliar chemistry and N cycling among forests of diverse history and composition. Ecology 83:339-355.

Oren, R., D. S. Ellsworth, K. H. Johnsen, N. Phillips, B. E. Ewers, C. Maier, K. V. R. Schäfer, H. McCarthy, G. Hendrey, S. G. McNulty, and G. G. Katul. 2001. Soil fertility limits carbon sequestration by forest ecosystems in a CO_2-enriched atmosphere. Nature 411:469-472.

Pfisterer, A. B., and B. Schmid. 2002. Diversity-dependent production can decrease the stability of ecosystem functioning. Nature 416:84-86.

Pielke, R. A., Sr. 2002. Overlooked issues in the U.S. national climate and IPCC assessments. Climatic Change 52:1-11.

Pinard, M. A., and J. Huffman. 1997. Fire resistance and bark properties of trees in a seasonally dry forest in eastern Bolivia. Journal of Tropical Ecology 13:727-740.

Prinn, R., H. Jacoby, A. Sokolov, C. Wang, X. Xiao, Z. Yang, R. Eckhaus, P. Stone, D. Ellerman, J. Melillo, J. Fitzmaurice, D. Kicklighter, G. Holian, and Y. Liu. 1999. Integrated global system model for climate policy assessment: feedbacks and sensitivity studies. Climatic Change 41:469-546.

Reich, P. B., J. Knops, D. Tilman, J. Craine, D. Ellsworth, M. Tjoelker, T. Lee, D. Wedin, S. Naeem, D. Bahauddin, G. Hendrey, S. Jose, K. Wrage, J. Goth, and W. Bengston. 2001. Plant diversity enhances ecosystem responses to elevated CO_2 and nitrogen deposition. Nature 410:809-812.

Reynolds, J. F. 2001. The role of controlled environment facilities in global change research. Pages 228-231 *in* H. A. Mooney and J. Canadell, editors. The Earth System: Biological and Ecological Dimensions of Global Environmental Change. Wiley, New York.

Ritchie, M. E., D. Tilman, and J. M. H. Knops. 1998. Herbivore effects on plant and nitrogen dynamics in oak savanna. Ecology 79:165-177.

Schellnhuber, H. J. 1999. 'Earth system' analysis and the second Copernican revolution. Nature 402 (Supp):C19-C23.

Schlesinger, W. H., J. A. Raikes, A. E. Hartley, and A. E. Cross. 1996. On the spatial pattern of soil nutrients in desert ecosystems. Ecology 77:364-374.

Shaver, G. R., L. C. Johnson, D. H. Cades, G. Murray, J. A. Laundre, E. B. Rastetter, K. J. Nadelhoffer, and A. E. Giblin. 1998. Biomass and CO_2 flux in wet sedge tundras: responses to nutrients, temperature, and light. Ecological Monographs 68:75-97.

Shaw, M. R., E. S. Zavaleta, N. R. Chiariello, E. E. Cleland, H. A. Mooney, and C. B. Field. 2002. Grassland responses to global environmental changes are suppressed by elevated CO_2. Science 298:1987-1990.

Siegert, F., G. Rücker, A. Hinrichs, and A. A. Hoffmann. 2001. Increased damage from fires in logged forests during droughts caused by El Niño. Nature 414:437-440.

Smith, S. D., T. E. Huxman, S. F. Zitzer, T. N. Charlet, D. C. Housman, J. S. Coleman, L. K. Fenstermaker, J. R. Seeman, and R. S. Nowak. 2000. Elevated CO_2 increases productivity and invasive species success in an arid ecosystem. Nature 408:79-82.

Smith, V. C., and M. A. Bradford. 2003. Do non-additive effects on decomposition in litter-mix experiments result from differences in resource quality between litters? Oikos 102:235-242.

Taylor, B. R., D. Parkinson, and W. F. J. Parsons. 1989. Nitrogen and lignin content as predictors of litter decay rates: a microcosm test. Ecology 70:97-104.

Tilman, D. 1989. Ecological experimentation: strengths and conceptual problems. Pages 136-157 *in* G. E. Likens, editor. Long-Term Studies in Ecology: Approaches and Alternatives. Springer-Verlag, New York.

Tilman, D., D. Wedin, and J. Knops. 1996. Productivity and sustainability influenced by biodiversity in grassland ecosystems. Nature 379:718-720.

Volk, M., P. A. Niklaus, and C. Körner. 2000. Soil moisture effects determine CO_2 responses of grassland species. Oecologia 125:380-388.

Volney, W. J. A., and R. A. Fleming. 2000. Climate change and impacts of boreal forest insects. Agriculture, Ecosystems & Environment 82:283-294.

Walker, B. H., and W. L. Steffen, editors. 1996. Global Change and Terrestrial Ecosystems. Cambridge University Press, New York.

Wardle, D. A., and K. S. Nicholson. 1996. Synergistic effects of grassland plant species on soil microbial biomass and activity: implications for ecosystem-level effects of enriched plant diversity. Functional Ecology 10:410-416.

Wardle, D. A., G. M. Barker, G. W. Yeates, K. I. Bonner, and A. Ghani. 2001. Introduced browsing mammals in New Zealand natural forests: aboveground and belowground consequences. Ecological Monographs 71:587-614.

Williams, D. L., P. Ineson, and P. A. Coward. 1999. Temporal variations in nitrous oxide fluxes from urine-affected grassland. Soil Biology & Biochemistry 31:779-788.

Zak, D. R., K. S. Pregitzer, P. S. Curtis, J. A. Teeri, R. Foge, and D. L. Randlett. 1993. Elevated atmospheric CO_2 and feedback between carbon and nitrogen cycles. Plant and Soil 151:105-117.

Zutter, B. R., G. R. Glover, R. J. Mitchell, and D. H. Gjerstad. 1997. Response of loblolly pine and sweetgum to intra- and inter-specific competition and influence of soil organic matter. Canadian Journal of Forest Research 27:2079-2087.