Multiscale Experiments in Coastal Ecology: Improving Realism and Advancing Theory

JOHN E. PETERSEN, W. MICHAEL KEMP, RICK BARTLESON, WALTER R. BOYNTON, CHUNG-CHI CHEN, JEFFREY C. CORNWELL, ROBERT H. GARDNER, DEBORAH C. HINKLE, EDWARD D. HOUDE, THOMAS C. MALONE, WILLIAM P. MOWITT, LAURA MURRAY, LAWRENCE P. SANFORD, J. COURT STEVENSON, KAREN L. SUNDBERG, AND STEVE E. SUTTLES

The Multiscale Experimental Ecosystem Research Center has conducted a series of mesocosm experiments to quantify the effects of scale—in terms of time, depth, radius, exchange rate, and ecological complexity—on biogeochemical processes and trophic dynamics in a variety of coastal habitats. The results indicate that scale effects can be categorized as (a) fundamental effects, which are evident in both natural and experimental ecosystems, and (b) artifacts of enclosure, which are solely attributable to the artificial environment in mesocosms. We conclude that multiscale experiments increase researchers' understanding of scale in nature and improve their ability to design scale-sensitive experiments, the results of which can be systematically compared with each other and extrapolated to nature.

Keywords: scale, mesocosm, experimental design, extrapolation

The recent literature in ecosystem ecology reveals two parallel trends. The first is an increased recognition of the importance of time and space scales as determinants of the patterns researchers observe and predict. This is evident in the steady increase in the number of journal articles listing *scale* as a keyword (figure 1a; Schneider 2001) and in the publication of numerous recent books devoted to scaling theory (e.g., Peterson and Parker 1998). The second trend is the increased use of manipulative field- and laboratory-based experiments to test ecological theory (figure 1b; Ives et al. 1996). Although scale and experimentation are conceptually linked, they remain awkwardly disconnected in practice; to date, experimental work in ecology has not taken full advantage of, or been effectively used to advance, scaling theory. Consequently, a debate has developed among ecologists regarding the appropriate scales for conducting ecosystem-level research.

It is commonly recognized that the scales selected for an investigation reflect a delicate balance among control (the ability to relate cause and effect, to replicate, and to repeat experiments), realism (the degree to which results accurately mimic nature), and generality (the breadth of different systems to which results are applicable; Kemp et al. 1980). Enclosed experimental ecosystems (micro- and mesocosms) have gained in popularity because they maximize opportunities for control (Odum 1984). This is particularly true in coastal aquatic habitats, where options for controlled field manipulation are constrained by complex bidirectional exchanges between land and sea. However, a number of critics have suggested that the high degree of control afforded by mesocosms comes with a concomitant reduction in realism (e.g., Schindler 1998). Others point out that problems of realism and scale are likewise evident in whole-ecosystem manipulations (Fee and Hecky 1992); both in mesocosms and in the field, researchers must extrapolate results from small

---

John E. Petersen (e-mail: john.petersen@oberlin.edu) is an assistant professor of environmental studies and biology at Oberlin College, Oberlin, OH 44074. W. Michael Kemp, Thomas C. Malone, Lawrence P. Sanford, and J. Court Stevenson are professors; Jeffrey C. Cornwell and Laura Murray are research associate professors; and Deborah C. Hinkle and Steve E. Suttles are senior faculty resident assistants at the University of Maryland Center for Environmental Science, Horn Point Laboratory, Cambridge, MD 21613. Rick Bartleson is a senior environmental scientist at the South Florida Water Management District Restoration Division, West Palm Beach, FL 33416. Walter R. Boynton and Edward D. Houde are professors at the University of Maryland Center for Environmental Science, Chesapeake Biological Laboratory, Solomons, MD 20688. William P. Mowitt is a staff biologist at Northwest Fisheries Science Center, Seattle, WA 98275. Chung-Chi Chen is an assistant professor in the Department of Biology at National Taiwan Normal University, Taipei 116, Taiwan. Robert H. Gardner is a professor at the University of Maryland Center for Environmental Science, Appalachian Laboratory, Frostburg, MD 21532. Karen L. Sundberg is a faculty research assistant in the Department of Biological Sciences at the University of South Carolina, Columbia, SC 29208.

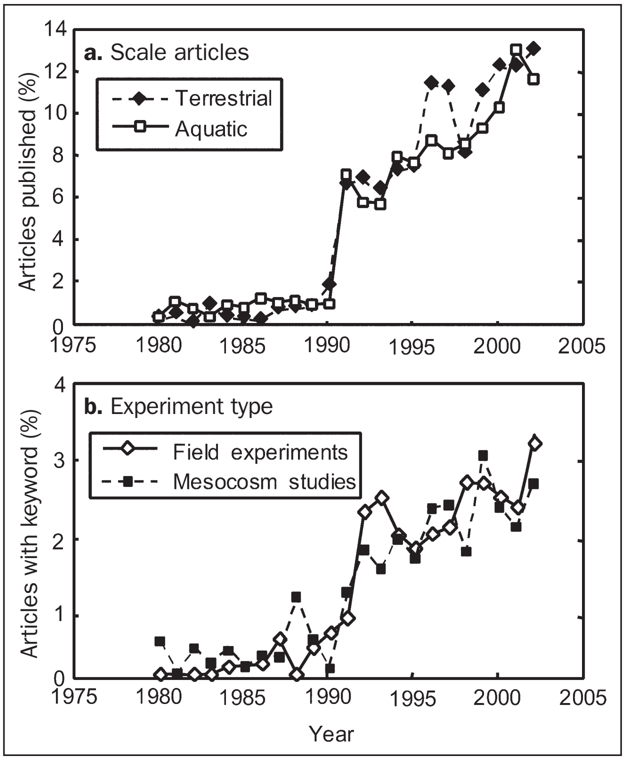


Figure 1. Trends in scaling concepts and experimental approaches in ecological studies between 1980 and 2002. (a) Separate searches were conducted by year for the term* scale *in keywords and abstracts of three journals emphasizing terrestrial research* (Ecology, Oikos, Oecologia) *and two journals publishing only aquatic research* (Limnology and Oceanography, Marine Ecology Progress Series)*. The number of articles identified in each year was standardized to the total number of articles published for that year in those journals and expressed in the graph as a percentage. (b) A similar search was conducted with the same five journals to identify field studies (operationally defined as those responding to the keyword* field experiment*) and mesocosm experiments (responding to the keywords* mesocosm, microcosm, enclosure, *and* limnocorral*). Given these operational definitions, it is likely that there was overlap between these two categories (e.g., mesocosm studies conducted in the field) and that most field and mesocosm studies were excluded because they did not use these keywords. As a result, the absolute number of papers and the actual balance between field and mesocosm studies may be in error; however, the temporal trends are representative (Kemp et al. 2001).

experimental systems to larger, deeper, more open, more biodiverse, and more heterogeneous ecosystems.

Parallel interests in scale and in manipulative experiments, coupled with concern regarding the relative merits of different approaches to experimentation, raise a series of questions that have been the focus of long-term studies at the Multiscale Experimental Ecosystem Research Center (MEERC) at the University of Maryland. Specifically, can scaling theory be used to improve the design of experiments so as to maximize realism? Conversely, can multiscale experiments be designed to test and advance scaling theory? Finally, can practical and empirical scaling rules be developed to facilitate comparisons among experiments and extrapolation of results from experiments to nature? The unique goal of MEERC has been to assess scale as an independent variable that drives ecological behavior. Over the last 10 years, we have conducted experiments in a variety of estuarine habitats to determine the effects of water depth, tank radius, ecosystem size, exchange rate, experimental duration, and ecological complexity on primary productivity, nutrient cycling, and trophic dynamics. Our goals in this article are to clarify "the problem of scale" (*sensu* Levin 1992) as it relates to manipulative research and to review MEERC studies and their implications for the use of mesocosms to understand nature.

The meaning of scale and its relevance for experiments

Discussions of ecological scale can be confusing, because the meaning of the word *scale* depends on its context (Schneider 2001). Among other things, *scale* can function as two parts of speech, a noun and a verb. As a noun, it refers to a set of variables that characterize the time, space, and complexity of organisms, ecosystems, and experiments. For instance, two key aspects of temporal scale that must be considered in experimental design are duration (length of experiment, organism generation time, perturbation response time) and frequency (timing of perturbations and sampling, exchange rates, variability in timing of inputs). Spatial scale is defined by variables such as length, depth, area, volume, shape, and heterogeneity (patchiness). Although a third scaling variable, ecological complexity (e.g., Frost et al. 1988), lacks fundamental dimensions of time and space, some have suggested that it can be used to characterize features such as species diversity, trophic dimensions (food-chain length, trophic diversity), environmental attributes (number of habitats, interconnections among habitats, biogeochemical diversity), and levels of ecological organization (population, community, ecosystem). Numerous researchers have found a correlation between scales of time and space, and sometimes between scales of time or space and scales of complexity, when they are plotted against each other (e.g., Schneider 2001). The verb form *to scale* refers to the act of relating or extrapolating information from one scale of time, space, or complexity to another. How can scientists extrapolate from small-scale experiments and observations to larger-scale natural ecosystems, from short-duration experiments to long time periods, from parts of nature to whole ecosystems, and from simplified ecosystems to the full complexity of nature? These questions should be of central concern to all ecological researchers.

Landscape ecologists emphasize two additional aspects of temporal and spatial scale, grain and extent. In the most general sense, *grain* refers to the smallest size or time scale that can be resolved. By analogy, it is useful to think of grain in terms of the pixels per square centimeter of film (spatial grain) and the frames per second shot by a movie camera (temporal grain). If grain is rough, viewers see a blocky, choppy moving image; if it is fine, they see a high-resolution image with smooth, continuous movement. In contrast to grain, *extent* is the total size and time scale of interest; it can be defined in terms of the size of the film (spatial extent; e.g., 16 versus 35 millimeters [mm]) and the total duration of the film (temporal extent). Three distinct contexts for grain and extent can be distinguished (Kemp et al. 2001), depending on whether the data of interest are observed in nature (observational scale), collected through experimental manipulations using a certain sample size and over a defined experimental area (experimental scale), or measured as intrinsic characteristics of a natural system, such as the school size of fish or the size of a watershed (natural or characteristic scale).

In all three contexts, grain and extent are dependent on the frame of reference and sampling technology used by the investigator and on the processes of interest (e.g., Wiens 1989). For instance, from the perspective of a population ecologist interested in a soil nematode, a small (e.g., 10 meters [m] x 10 m) plot and a specific season (e.g., summer) may define the scale of experimental extent. From the perspective of an ecosystem scientist tracing watershed biogeochemistry, on the other hand, this same small plot size and single-season duration may represent experimental grain, and the size of the entire watershed and a period of several years may appropriately define experimental and characteristic extent. From the perspective of a global change scientist using remote sensing technology, the watershed may represent experimental and characteristic grain size (i.e., a single pixel) in a model that defines extent as the regional landscape or even the whole biosphere over decades or centuries. Theory suggests that it is challenging to make inferences about dynamics that operate at scales finer than observational grain size or broader than observational extent (Wiens 1989). Ecologists' need to extrapolate beyond the extent scale of their manipulative experiments (i.e., to the larger scales of nature) therefore represents a significant challenge.

Several scaling concerns must be addressed when using mesocosm results to predict effects in natural aquatic ecosystems. The first and most obvious is that experimental systems are constrained in extent. With a median duration of 49 days and a median volume of 1.7 cubic meters (m^3) (Petersen et al. 1999), aquatic mesocosm studies are obviously brief and small relative to the natural grain and extent scales that characterize many important ecological processes. A second problem is the presence of walls, which restrict biological, material, and energy exchange with the outside world and provide a substrate for growth of undesirable (but potentially influential) organisms on this artificial edge habitat. A third problem is that a host of experimental design decisions—such as how many replicates to include per treatment and whether to control light, mixing, and other properties—tend to be constrained by and vary together with choices of size, duration, and complexity (figure 2). Finally, the relative importance of the air–water area, sediment–water area, and wall area, in relation to each other and to water and sediment volume, changes with the physical dimensions of a mesocosm.

An analysis of aquatic studies conducted in cylindrical planktonic–benthic mesocosms revealed that experimenters gravitate toward a depth-to-radius ratio of approximately 4.5 (figure 3; Petersen et al. 1999). As a consequence of this

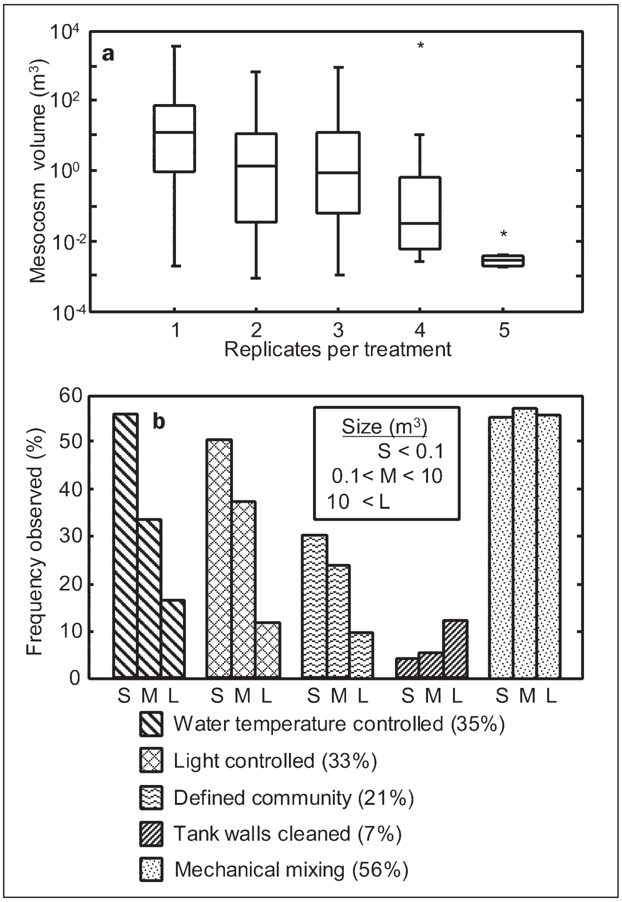


Figure 2. (a) Plot of mesocosm volume versus number of replicates per treatment. Median values are represented by the bar within a box, and the 75th and 25th percentiles (i.e., the interquartile range) by the top and bottom of the box. The ends of the "whiskers" represent the farthest data point within a span that extends 1.5 times the interquartile range from the 75th and 25th percentiles. Data outside this span are graphed with asterisks. (b) Relationship between mesocosm size and the presence of various design characteristics in a quantitative review of the mesocosm literature. Size categories (small, medium, or large, in cubic meters) are indicated in the legend. The y-axis represents the percentage of articles in a given size class for which the design characteristic indicated is present. The overall percentage of experiments for which a given characteristic is present is indicated in parentheses within the key. "Defined community" indicates that individual populations were selectively added to create the mesocosm community.

bias, in general, larger mesocosms simultaneously are less influenced by wall artifacts, have less sediment area per unit volume, and have less surface area available for gas and light exchange per unit volume than do smaller systems (figure 3b, 3c). Collectively, these scaling problems complicate interpretation, comparison, and extrapolation of findings from mesocosm experiments. Unfortunately, parallel scaling problems also exist for field experiments. For example, replication tends to decrease with increasing plot size (e.g., Kareiva and Andersen 1988), and experimental lakes and field plots tend to be orders of magnitude smaller and shallower than the natural systems for which inferences are drawn (e.g., Fee and Hecky 1992).

One might conclude from the preceding discussion that reductions, artifacts, covariation, and distortions in scale pose an almost insurmountable obstacle to designing experimental ecosystem studies. Alternatively, we have chosen to view these problems as interesting research opportunities that can be used to advance theoretical and practical understanding of the "science of scale" *(sensu* Meentemeyer and Box 1987). Specifically, MEERC studies were designed to shed light on two classes of effect: (1) fundamental effects of scale, which are evident in both natural and experimental ecosystems (e.g., the effects of water depth), and (2) artifacts of enclosure, which are attributable to the artificial environment in experimental ecosystems (e.g., the effects of wall growth). To this end, we have conducted a series of experiments in which we have systematically manipulated time, space, and complexity in a variety of estuarine habitat types (figure 4, table 1). In the sections below, we draw on these experiments to address the following scale-related questions:

- What are the effects of depth and tank radius on the dynamics of experimental planktonic ecosystems?
- To what extent can depth-scale effects observed in these mesocosm studies be extrapolated to nature?
- How do trophic depth and species diversity (two features of ecosystem complexity) affect estuarine ecosystem dynamics?
- How does the rate at which water is exchanged within estuarine habitats and among different types of estuarine habitats affect ecological dynamics?
- To what extent does physical scale affect ecological variability?
- What rules and tools are available for improving the design of mesocosm experiments so that they more realistically represent the dynamics of larger-scale natural ecosystems?

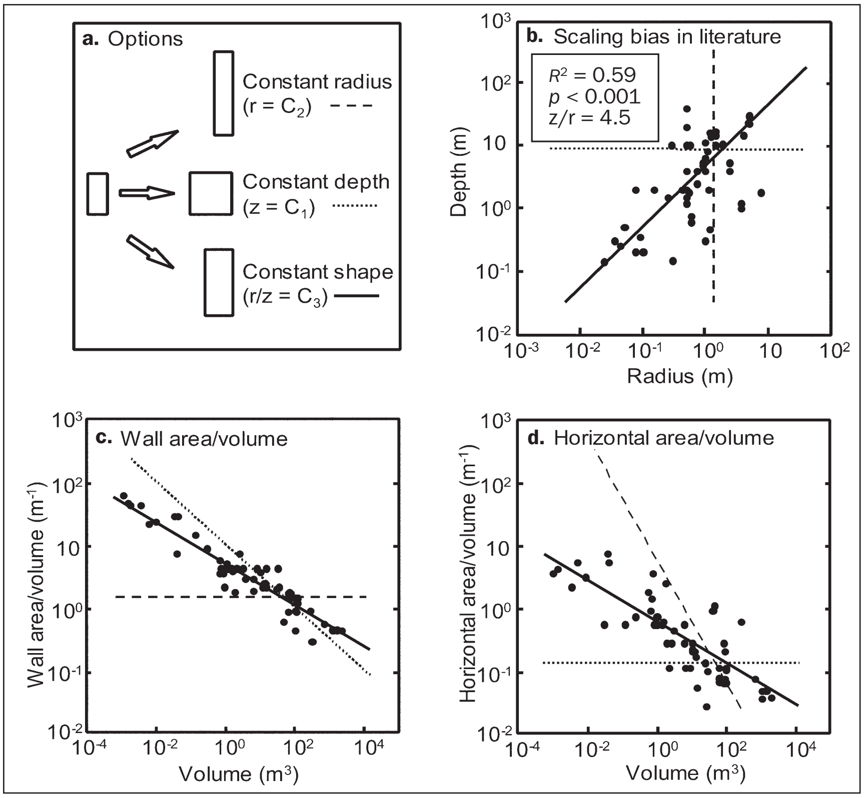


Figure 3. (a) Three options exist for holding scaling relationships constant as mesocosm size is increased: Radius can be held constant, depth can be held constant, or shape (the ratio of depth to radius) can be held constant. (b) Plots of radius versus depth (in meters) for cylindrical mesocosms reveal a bias toward scaling for constant shape. As a result of this bias, (c) wall area per unit volume (in cubic meters) and (d) horizontal surface area (e.g., water surface and benthic area) per unit volume simultaneously decrease with increasing mesocosm size. Dotted lines represent scaling for constant depth and are placed at values corresponding with median depth. Dashed lines represent scaling for constant radius and are placed at median radius. Solid lines represent scaling for constant shape and are derived from linear regression of radius (r) versus depth (z). Source: Petersen et al. 1999.

Experiments to determine the effects of physical dimensions

The first priority of MEERC was to assess the effects of mesocosm size and shape on the dynamics of planktonic–benthic ecosystems (figure 4a). Fifteen cylindrical mesocosms were constructed, with five distinct dimensions, three volumes, and three replicates per dimension (figure 5). These mesocosms were organized into three series: one with a constant depth (A, C, and E tanks; depth = 1.0 m), one with a constant shape (B, C, and D tanks; radius-to-depth ratio = 0.57), and one with constant volume (D and E tanks, volume = 10 m^3; A and B tanks, volume = 0.1 m^3). Mesocosms received artificial light on a 12:12 light–dark cycle and an exchange of filtered estuarine water at a rate of 10% per day; they were mixed to mimic turbulence in a tidal environment

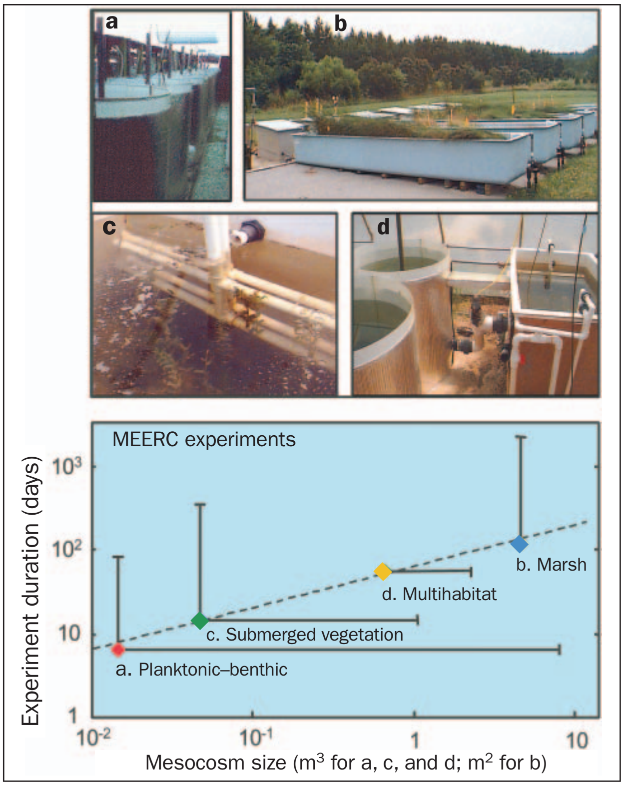
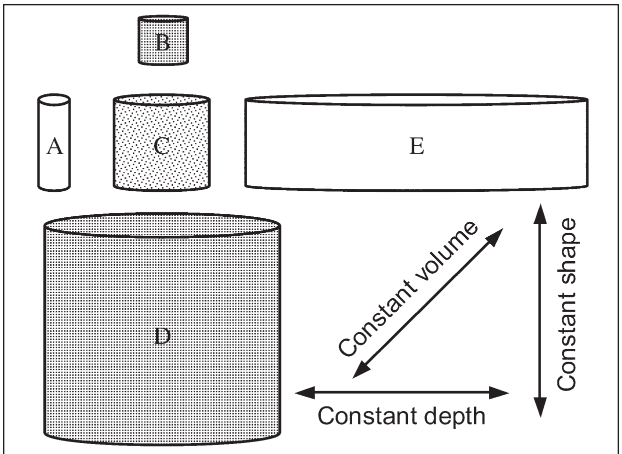


Figure 4. Multiscale, multihabitat experimental ecosystems: (a) planktonic–benthic systems, (b) marsh, (c) submerged aquatic vegetation, and (d) linked multihabitat systems. Lower panel shows the range of scaling dimensions in experiments conducted in these different experimental systems (letters are used for the same habitats in the two panels). For each habitat type, diamond-shaped points represent the smallest mesocosm and the shortest-duration experiment. Bars extend out to the largest experiment and the longest-duration experiment (there is no vertical bar for the multihabitat systems because all the experiments in these systems were 45 days in duration). Only one size of marsh mesocosm was used in the studies. Photographs: John Petersen.

and initiated with unfiltered water and bottom sediment (10 centimeters deep, homogenized) from the Chesapeake Bay (Petersen et al. 1997). It should be noted that although environmental conditions were carefully controlled to simulate critical features of the estuarine environment, this design nevertheless represents an intentional simplification of the physical and biological complexity present in nature. The experimental extent of even the largest system in this study was smaller than plankton patch size (i.e., smaller than characteristic grain) in nature.

A host of variables were measured (table 1). Primary productivity was used as an integrated measure of dynamic ecological responses to variations in system size and shape. A set of simple scaling hypotheses was developed based on differences in wall area and water-column depth. Depth-scaling hypotheses started with the understanding that primary productivity in temperate coastal ecosystems often experiences a seasonal shift from light-limitation in the spring to nutrient-limitation in the summer. An important dimensional difference between these two limiting factors is that light energy is received on an areal basis (e.g., units of micromoles [μmol] per square meter [m^2] per second) and is then absorbed as it travels down to deeper parts of the water column. In contrast, plankton experience nutrients on a volumetric basis (e.g., moles per m^3), and concentration is relatively constant over depth in a well-mixed water column.

Depth-scale hypotheses. These dimensional differences in nutrients and light suggest two simple depth-related scaling hypotheses that were tested in our mesocosm experiments. First, because light is experienced on an areal basis, under purely light-limited conditions gross primary productivity (GPP, or net daytime molecular oxygen (O_2) production + nighttime O_2 respiration) might be expected to be constant among different depth systems when expressed on an areal basis ($GPP_{area} = C_1$, where C_1 = a constant; units = grams (g) O_2 per m^2 per hour). If this is the case, then by definition, GPP expressed per unit volume must be inversely proportional to depth: $GPP_{vol} = C_1/z$ (z = depth of mesocosm or mixed layer, in m; GPP_{vol} units = milligrams O_2 per m^3 per hour). In contrast, because phytoplankton contact nutrients on a



Tank	Radius (m)	Depth (m)	Volume (m^3)
A	0.18	1.00	0.1
B	0.26	0.46	0.1
C	0.57	1.00	1
D	1.22	2.15	10
E	1.78	1.00	10

Figure 5. Relative sizes and shapes of cylindrical experimental ecosystems viewed from the side in size–shape effect experiments. Constant-depth series (A, C, E; depth = 1.0 meter) and constant-shape series (B, C, D; radius:depth ratio = 0.57) intersect at the intermediate 1.0-cubic-meter mesocosms (C). Three replicates of each mesocosm type were used in experiments.

Table 1. Summary of scale experiments conducted at the Multiscale Experimental Ecosystem Research Center.

Habitat	Scaling variable manipulated	Primary dependent variable	Reference
Planktonic–benthic	Size and shape	Primary productivity	Petersen et al. 1997
	Size and shape	Relative importance of wall growth	Chen et al. 1997
	Size and shape	Bacterioplankton dynamics	Sanford et al. 2001
	Size and shape	Microzooplankton dynamics	Merrell 1996
	Presence of wall growth	Primary productivity, nutrient dynamics	Chen et al. 2000
	Mixing intensity	Primary productivity, community dynamics	Petersen et al. 1998
	Mixing intensity	Gypsum dissolution	Porter et al. 2000
	Mixing intensity	Boundary layer flow	Crawford and Sanford 2001
	Water exchange rate	Primary productivity, community dynamics	
	Light intensity	Primary productivity	
	Coupling between habitats	Plankton and biogeochemistry	Porter 1999
	Trophic complexity	Fish growth, trophic dynamics	Mowitt 1999
	Continuous versus pulsed predation	Fish growth, trophic dynamics	Mowitt 1999
Submerged aquatic vegetation	Nutrient exchange	Plant growth, competition between plants and epiphytes	Sturgis and Murray 1997, Murray et al. 2000
Marsh	Species diversity, groundwater nutrient concentration	Biomass, nutrient retention	
	Disturbance	Regrowth following fire	Schmitz 2000
Multiple-habitat mesocosms	Degree of coupling in linked planktonic–benthic and submerged aquatic vegetation systems		

Note: Experiments without references represent the unpublished work of one or more coauthors of this article.

volumetric basis, under purely nutrient-limited conditions, primary productivity should be constant when expressed per unit volume: $GPP_{vol} = C_2$. In this nutrient-limited case, by definition, productivity expressed per unit area should be directly proportional to depth: $GPP_{area} = C_2 \times z$.

Radius-scale hypothesis. A review of the literature revealed that periphyton growth is cleaned from tank walls in only a small percentage of mesocosm studies (figure 2b). We therefore conducted experiments with tanks of different radius to quantify the artifacts of scale associated with this wall growth. A reasonable starting assumption in such experiments is that the relative contribution of wall productivity to total system productivity should be proportional to the ratio of wall area to water-column volume. In a cylindrical mesocosm, the wall area-to-volume ratio is inversely proportional to the radius ($[2\pi r \times z]/[\pi r^2 \times z] = 2/r$, where r = radius). Based on these basic geometric considerations, a simple hypothesis for the relative contribution of walls to total gross primary productivity among systems of different radius is that $GPP_{wall}/GPP_{total} = C_3/r$ (C_3 = a constant).

To test these depth and radius hypotheses, 4- to 6-week experiments were conducted in the MEERC facility under both high and low nutrient conditions. Special incubation chambers with internal mixing were constructed to allow wall and water column productivity to be measured independently (Petersen and Chen 1999). Data from these experiments (figure 6a, 6b) were consistent with both of our depth-scale hypotheses: Productivity in the different depth tanks was constant when expressed per unit area under light-limited conditions and was constant when expressed per unit volume under nutrient-limited conditions. As an additional test of these concepts, we applied a pulse of nutrients to the ecosystems during the low-nutrient experiment. Consistent with our hypothesis, this resulted in a temporary shift away from constant productivity per unit volume toward constant productivity per unit area (figure 6c). The finding that depth can regulate the balance between total system production and respiration has important implications for trophic relations and biochemistry in natural ecosystems (Fee 1979, Petersen et al. 1997). In a more general sense, the results of these experiments suggest that hypotheses regarding the fundamental effects of scale, which would be difficult to test in natural ecosystems, can be successfully tested using multiscale mesocosm experiments.

Results from this series of experiments were also generally consistent with our expectation that the relative contribution of wall periphyton to total system productivity and associated nutrient uptake would decrease with increasing radius (Chen et al. 1997, 2000). The results do not, however, fall cleanly along the line predicted by the radius-scale hypothesis (figure 7a). Indeed, when wall productivity is expressed per unit of wall surface area, periphyton growth rate on tank walls is a hyperbolic function of tank radius (figure 7b). This indicates that scale effects can be more complicated than simple geometric considerations suggest. For example, in our experiments, it is possible that periphyton access to light or to nutrients increased with system width. In a general sense, these results indicate that multiscale experiments may sometimes reveal unanticipated but empirically quantifiable effects of scale.

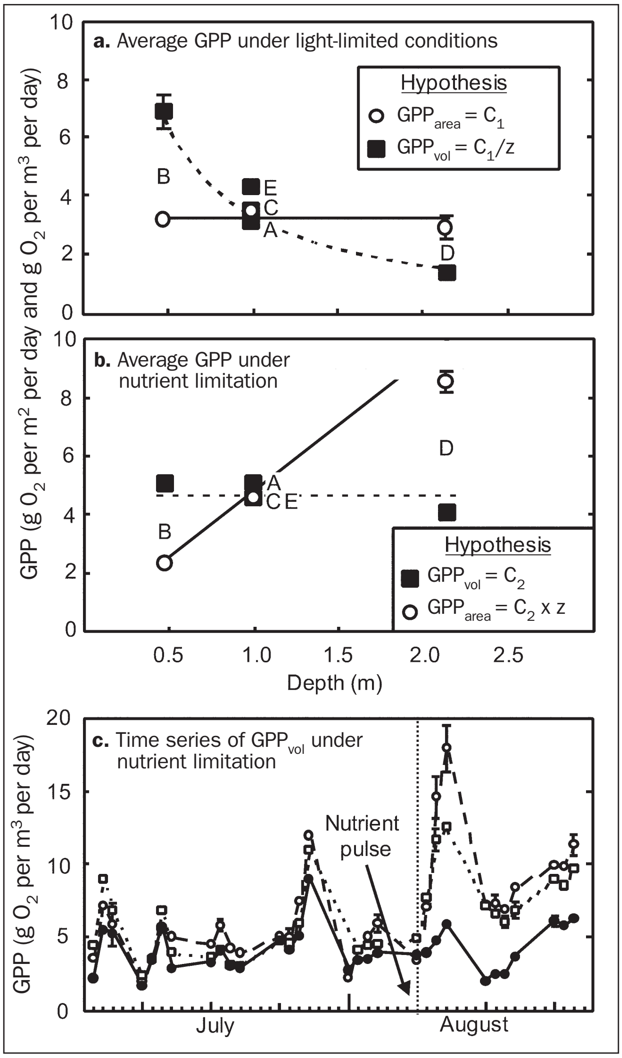


Figure 6. Gross primary productivity (GPP), in grams of oxygen per unit volume per day (filled squares) and per unit area per day (unfilled circles), averaged over the prenutrient pulse period for each dimension of mesocosm under (a) light-limited and (b) nutrient-limited conditions. Dashed lines represent hypothesized relationship between GPP per unit volume (GPP_{vol}) and depth (z) in meters; solid lines represent hypothesized relationship between GPP per unit area (GPP_{area}) and z. C_1 and C_2 are constants for light- and nutrient-limited conditions, respectively. (c) Time series of GPP_{vol} data from the nutrient-limited experiment. A pulse of nutrients was added to the system at the point indicated by the dashed vertical line. In all three panels, error bars are standard error of the mean. Bars are excluded in cases where the error is smaller than the diameter of the symbols. Source: Petersen et al. 1997.

Extrapolating effects from mesocosms to nature

The experiments discussed above illustrate the potential for quantifying scaling relationships through multiscale experiments. However, we recognize that, as in most ecological experiments, the experimental extent scale of even MEERC's largest experimental systems is considerably smaller than the natural grain and extent scales of the natural ecosystems to which we wish to extrapolate results. Scaling theory predicts that certain patterns and processes become evident only as scale is increased beyond certain thresholds of extent (Wiens 1989). Furthermore, our own hypotheses and data indicate that scaling patterns tend to be nonlinear (e.g., figures 6 and 7). So it is possible, for instance, that patterns that are scale dependent in our mesocosm experiments become scale independent at larger scales of natural systems (figure 8a, solid line). Likewise, it is possible that relationships that we see as scale independent in mesocosms are functions of scale in larger natural ecosystems (figure 8a, dotted line). Finally, it is possible that thresholds exist over which small changes in scale result in dramatic and possibly

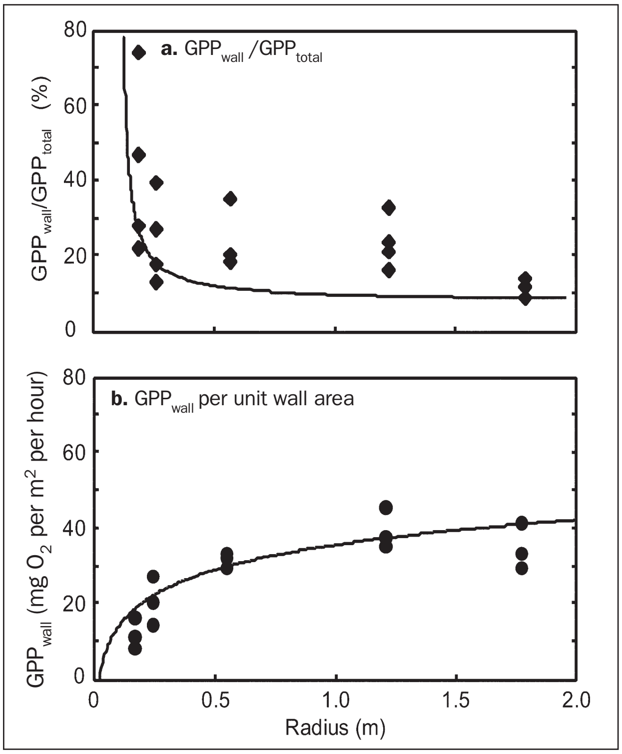


Figure 7. (a) Relative contribution of wall periphyton to the total gross primary productivity (GPP), in milligrams of oxygen per square meter per hour, within mesocosms of different dimensions. The curved line is a least-squares fit of the hypothesized inverse relationship between relative productivity and tank radius. (b) Productivity of the wall periphyton expressed per unit of wall area in the same mesocosms. The curved line is a best-fit line.

discontinuous changes in ecological dynamics. Given these possibilities, it is important that findings from multiscale experiments be validated with data collected from a range of larger-scale ecosystems in nature.

The MEERC mesocosms were derived from, and intended to serve as models of, Chesapeake Bay ecosystems. Therefore, it seems reasonable to compare our results with data from parallel field studies conducted in the mainstem Chesapeake Bay and in its tributaries (figure 8b, 8c, 8d). To assess the legitimacy of extrapolating the results from our experiments to natural systems, we calculated a least-squares fit for the general equation describing our depth-scale hypotheses ($Y = C_1/z$, where Y = a depth-dependent ecological variable; figure 6a). We used data from light-limited (spring) experiments both for mesocosms alone (broad gray lines in figure 8b, 8c, 8d) and for mesocosms plus natural ecosystems (solid black lines in figure 8b, 8c). In one case, this was done with GPP used as the dependent variable (figure 8b). In the second case (figure 8c), we assumed that zooplankton were limited by primary productivity rather than by predation, and therefore we used a similar equation to express zooplankton biomass (ZPB) as an inverse function of depth ($ZPB_{vol} = B_1/z$, where z = water depth; ZPB_{vol} units = g carbon per m^3).

For both GPP and ZPB, coefficients derived from mesocosm data alone were almost identical to those derived by combining data from mesocosms and natural ecosystems (figure 8b, 8c). It is worth noting, however, that GPP data for the Patuxent estuary fall some distance from the fitted equation. Had we included data for other natural ecosystems, we would anticipate considerable scatter around the regression, because many factors other than depth vary among estuaries. Indeed, in many ways, this is the point: The great strength of multiscale mesocosm experiments is that these other factors, such as nutrient loading, can be held constant so that scaling effects can be successfully isolated. This comparison suggests that, at least for some variables and under some circumstances, scaling relationships revealed in mesocosms are robust and can be directly extrapolated to natural systems.

For other variables, however, the potential value of mesocosms for revealing fundamental scaling relationships and the potential for direct extrapolation may be more complicated. For instance, we anticipated that benthic nitrogen recycling would also follow our depth-scale hypothesis. However, in this case, although an inverse equation neatly fits data from our mesocosm experiments (gray line in figure 8d), it is abundantly evident that this equation does a poor job of predicting data gathered from larger natural ecosystems (dotted line in figure 8d). There are a number of explanations for why scaling relationships for nitrogen regeneration might differ between mesocosms and nature. For example, in the mesocosms in our studies, the sediment was mixed immediately before each experiment, and this disturbance may have accentuated initial regeneration rates. In addition, although the horizontally rotated impellers used to mix our experimental systems produced realistic mixing in the water column, these impellers produced much less realistic mixing at the sediment–water interface. The lesson is that although multiscale experiments can, in some cases, be used to identify valid scaling relationships, unrealistic biological and physical conditions within mesocosm studies (i.e., an inadequate representation of physical complexity) can distort ecological dynamics and result in erroneous conclusions.

Effects of ecosystem complexity

A rule of thumb in any type of modeling is that the best model contains the minimum degree of complexity necessary to accurately capture the dynamics under investigation. Our results suggest that a relatively simple planktonic experimental ecosystem may be sufficient to capture a range of depth-related scaling patterns evident in deeper natural planktonic ecosystems. But many important coastal dynamics are determined by more complex interactions within and among trophic levels and among habitats that were excluded from the MEERC planktonic–benthic mesocosms. One of our goals, therefore, was to explore the effects of ecological complexity by varying both trophic and habitat complexity (figure 4, table 1).

Adding trophic complexity to planktonic–benthic ecosystems.

Zooplanktivorous fish were added to planktonic–benthic mesocosms in order to (a) determine how the size and shape of experimental ecosystems affect fish behavior and (b) explore the effects of greater trophic complexity on ecosystem dynamics. Our scaling hypotheses for the initial planktonic experiments assumed bottom-up control of the phytoplankton by light or nutrients. Adding fish to these systems allowed for top-down control through predation on zooplankton.

In two experiments, bay anchovies *(Anchoa mitchilli),* a small, schooling zooplanktivorous fish, were added in equal density to the C, D, and E tanks (figure 5). Bay anchovies are the most abundant fish in Chesapeake Bay and constitute a major food source for predatory fish (Hartman and Brandt 1995). In a 30-day experiment, anchovies were stocked at densities of two individuals per m^3, and in a 74-day experiment, they were stocked at densities of four individuals per m^3. Mesocosms that contained no fish were also included as controls. Growth rates were determined by weighing fish at the start and end of each experiment. Zooplankton abundance, chlorophyll *a,* primary productivity, and nutrient concentration were measured at 3- to 7-day intervals.

As observed in many previous studies (e.g., Carpenter et al. 1985), the presence of zooplanktivorous fish did indeed reduce the number, biomass, and size of zooplankton (Mowitt 1999). Unique to this study, however, were the dramatic effects of mesocosm size, shape, and experimental duration on fish growth rates (figure 9). In the shorter (30-day) experiment, fish growth was linearly related to the ratio of volume to wall area. The results of an individual-based simulation model constructed in parallel with our experiments suggest

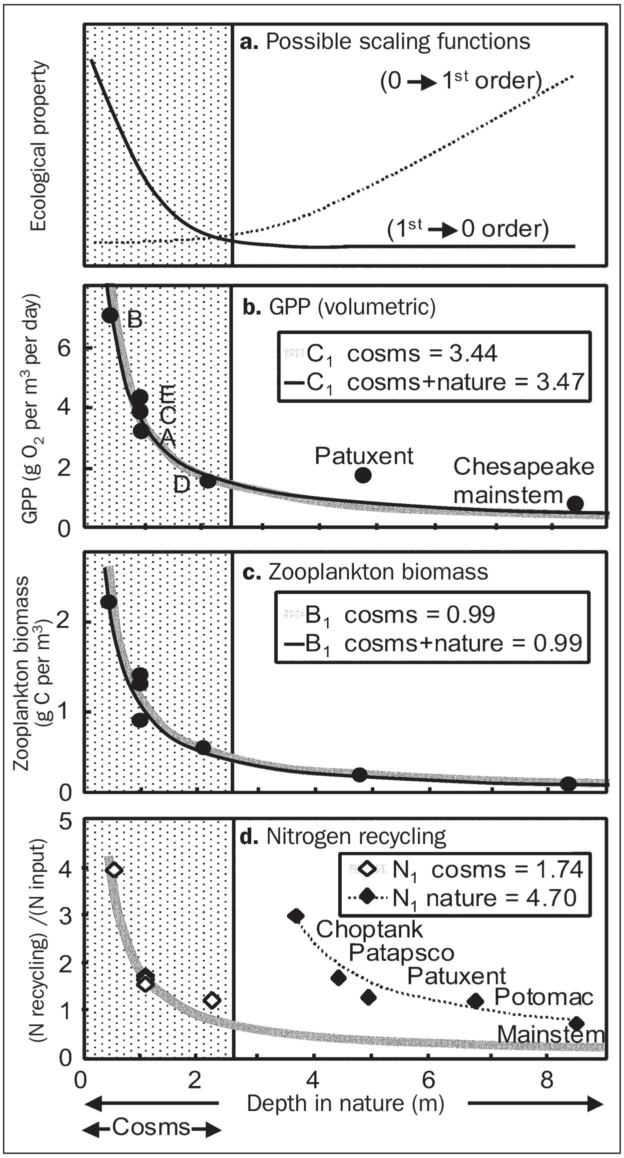


Figure 8. Extrapolating and verifying scaling relationships from mesocosms to nature. (a) Hypothetical responses of two distinct ecological properties with changes in scale in the size range of mesocosms (shaded region of graph) and nature. Trajectories shown indicate how different properties may be affected differently by changes in scale. (b) A case in which scaling patterns found in mesocosms can be directly extrapolated to nature: gross primary productivity (GPP) for the upper mixed layer of the Chesapeake Bay, in grams of oxygen per cubic meter per day, taken from measurements of oxygen change in light and dark bottles. Gray lines are regressions based on mesocosm data alone; black lines are regressions that include points from nature. (c) Direct extrapolation of scaling patterns from mesocosms to nature is also possible for zooplankton biomass, measured in grams of carbon per cubic meter. Zooplankton field data are for adult and nauplii copepods from oblique tows in the Chesapeake Bay. (d) The scaling relationship for nitrogen recycling in the Chesapeake Bay, on the other hand, cannot be directly extrapolated from mesocosms to nature; the patterns look similar to those of GPP and zooplankton biomass, but the magnitudes and fitting equations differ for data from mesocosms and from nature. The constants C_1, B_1, and N_1 are fitted regression coefficients and correspond with the equation discussed in the section of the article on depth-scale hypotheses. Sediment–water fluxes of dissolved inorganic nitrogen are measured from intact cores. Data for all three Chesapeake Bay examples are for the springtime mean. Source: Brownlee and Jacobs 1987, Smith and Kemp 1995, Kemp et al. 2001.

that this scale dependence in fish growth is largely attributable to negative effects of wall encounters on foraging efficiency (Heath and Houde 2001). In the 30-day experiment, fish growth rates in the two 10-m^3 systems fell within the range observed for anchovy of similar size (50 mm) in Chesapeake Bay (figure 9; Newberger and Houde 1995). Although growth rates in a longer (74-day) experiment also increased with the ratio of volume to wall area, these rates were much lower than those in the 30-day experiment. We surmise that fish growth became food-limited during the second half of the 74-day experiment because zooplankton populations were unable to compensate for the continuous grazing pressure of fish. In this experiment, then, the effect of increased experimental duration (i.e., the temporal extent of the experiment) also represented an artifact of reduced system size (the spatial extent of the experiment). In summary, we found a definable window for experimental extent (size $\geq$ 10 m^3, duration $\leq$ 30 days) in which our experimental ecosystems produced realistic fish behavior. The dimensions of this research window obviously depend on the organism and habitat under investigation, but our results suggest that the dimensions necessary to achieve realistic behavior can be experimentally determined.

Varying species diversity in experimental marsh ecosystems. The debate over the relationship between ecological function and various attributes of biological diversity has been fueled in large part by the results of experiments conducted in terrestrial and microbial mesocosms and in relatively small research plots (e.g., Tilman et al. 1997, Naeem and Li 1998). Because tidal estuarine marshes have long been postulated to play an important role in processing nutrients and organic matter at the land–sea interface, we designed an experiment to evaluate the effects of species diversity on this function. Six marsh mesocosms were constructed using fiberglass tanks (6 m long, 1 m wide, 1 m deep) angled at a 20:1 slope (figure 4b). Continuous groundwater flow was delivered with high and low nutrient concentrations at rates designed to match average monthly input to local coastal marshes (Staver and Brinsfield 1993). In addition, pumps were employed to simulate a 12.5-hour tidal exchange with filtered water from an adjacent estuary. The experimental systems were initiated with a combination of homogenized sediments and

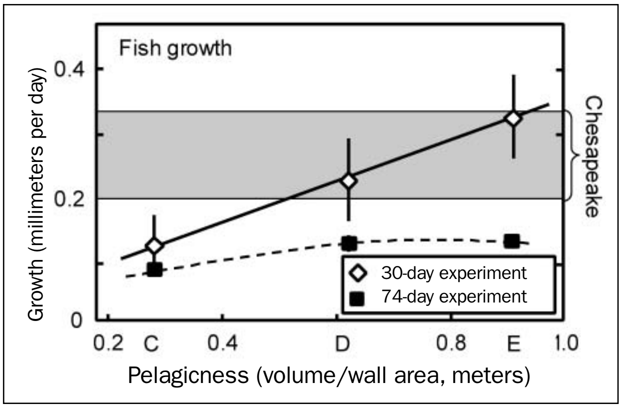


Figure 9. Fish growth rates in 30- and 74-day experiments in which fish were added in equal density (number of fish per unit volume) to different-sized cylindrical mesocosms. The shaded region represents the range of growth rates found for fish raised in Chesapeake Bay (Newberger and Houde 1995). Error bars are the standard error of the mean. Source: Mowitt 1999.

sediment plugs containing marsh plants taken from nearby natural habitat. A low-diversity treatment was planted and weeded regularly to maintain three common plant species found in Chesapeake marshes, *Spartina alterniflora, Spartina patens,* and *Distichlis spicata,* and no macrofauna were intentionally added. A high-diversity treatment was initiated with the same three plants plus six additional plant species and eight animal species. The experiment was run for 5 years to evaluate long-term dynamics. We should note that while we were able to preserve the plant species diversity over the temporal extent of the experiment, we were unsuccessful in maintaining the additional animals introduced in the high-diversity treatment.

Both low- and high-diversity marshes proved to be effective at removing nutrients from groundwater for the duration of the experiment. This supports the long-held view that coastal marshes play a critical functional role in eutrophic landscapes (Odum 1961). However, our most striking finding was that plant diversity had no significant effect on any of the functional variables examined during the experiment. The fact that others have observed significant effects of diversity on functional attributes such as productivity in terrestrial environments (e.g., Symstad et al. 2003) implies that plant diversity may be more important in some types of ecosystems than in others. In a larger sense, these results should serve as a caution against direct extrapolation of findings among ecosystem types or from generalized experimental ecosystems that are designed to represent broad classes of ecosystems to specific ecosystem types.

Although diversity had no statistically detectable effect on ecosystem function, the temporal extent of the experiment had a substantial effect on productivity in our marsh mesocosms. During the first 2 years of the study, plant growth was considerably higher than the typical values reported for natural marshes. We speculate that this is attributable to ample nutrients in the disturbed sediments, to light penetration, and to available ground surface for colonization early in the experiment. During subsequent years, growth rates fell to levels within the reported ranges for plants in natural marshes. Most marsh mesocosm experiments are of short duration (typically days to weeks; Petersen et al. 1999). The rapid initial growth evident in our experiment early on implies that the treatment effects observed in such short-duration experiments may be confounded by artifacts of scale associated with experimental extent scales that are far shorter than the natural extent scales that characterize the life cycles of marsh plants.

Effects of water exchange

The rate of water exchange with surrounding ecosystems is a scaling characteristic that controls many important processes in aquatic systems. Estuarine ecosystems frequently exhibit high rates of primary and secondary productivity, and this is often attributed to large material exchange resulting from the estuaries' position at the interface between the watershed and the tidal ocean (e.g., Odum 1961). Although exchange incorporates both temporal and spatial scale, it is often convenient to express water exchange in terms of residence time (i.e., time required for incoming water to replace the entire volume of the basin or container) or, alternatively, as exchange rate (1/residence time). Flow-through chemostat experiments are commonly used to study phytoplankton growth; however, few ecosystem-level studies have attempted to simulate exchange rates that characterize specific natural ecosystems, and fewer still have explicitly assessed the effects of different exchange rates on ecological dynamics. Indeed, most mesocosm experiments are closed to water exchange (Petersen et al. 1999), a situation that almost certainly distorts ecological dynamics. To examine the effects of exchange, we conducted experiments in and among a variety of estuarine habitats.

Water exchange in planktonic–benthic systems. We used nine 1-m^3 mesocosms (the C tanks in figures 4a and 5) to determine the effect of different exchange rates (0%, 10%, and 30% exchange of filtered estuarine water per day) on plankton dynamics. In 4-week experiments conducted with high-nutrient influent waters typical of spring conditions, primary productivity and phytoplankton biomass increased with increasing exchange rate (figure 10a), while zooplankton biomass decreased with increasing exchange (figure 10b). When the same experiment was repeated using low-nutrient influent waters (summer), the pattern of increase in productivity and phytoplankton biomass was similar (figure 10a). In this case, however, zooplankton abundance increased with increasing exchange rate (figure 10b). Simulation modeling studies corroborated the conclusion that a transition in the effects of exchange occurred under these two scenarios: In the high-nutrient experiment, the primary producers were controlled from the top down by zooplankton grazing. Zooplankton abundance, in turn, was controlled by the rate at which these organisms washed out of the mesocosm. In the low-nutrient experiment, phytoplankton and zooplankton were both controlled from the bottom up by the availability of nutrients. Thus, we found that the consequences of changing water exchange rates on a planktonic community depend on the characteristics of incoming water.

Water exchange and nutrient competition between submerged aquatic plants and their epiphytes. Several mesocosm experiments have demonstrated that nutrient enrichment can inhibit the growth of sea grasses and other submersed plants by stimulating growth of (and associated shading by) epiphytic algae (e.g., Short et al. 1995). Other studies, however, have reported little response of epiphytes and sea grasses to nutrient enrichment (e.g., Taylor et al. 1995). A close examination of the literature revealed that these discrepancies in observed responses might be related to differences in water exchange rates (Murray et al. 2000). Hence, we conducted a series of 6-week experiments in 1-m^3 mesocosms to examine effects of exchange rate on *Potamogeton perfoliatus* and associated epiphytes (Murray et al. 2000). Tanks with four exchange rates (100%, 300%, 600%, and 1200% per day; three replicates per treatment) were planted with submersed plants. Separate experiments were conducted with low concentrations and high concentrations of nutrients in replacement water (2 to 4 and 20 to 24 μmol ammonium, with a 10:1 molar ratio of nitrogen to phosphorus).

We found that relative responses of plants and epiphytes to nutrient concentrations were indeed strongly influenced by exchange (figure 10c, 10d). At low water exchange rates, epiphyte growth was inhibited because host plants were able to deplete nutrients from the water column, while at higher exchange rates, epiphytes flourished because host plant uptake had little effect on nutrient concentrations. A simulation model that was developed in parallel corroborated this interpretation and also revealed the importance of initial plant biomass and timing of nutrient additions during the life cycle of the host plant (Murray et al. 2000).

Effects of exchange among habitats. Waves, tides, and river flow induce the exchange of organisms, materials, and energy among estuarine habitats. Although this exchange almost certainly affects dynamics within individual habitats, it has proved challenging to include multiple habitats within a single mesocosm in a way that retains realistic proximities, interactions, and relationships. This has led some researchers to design coupled mesocosms that move water through a

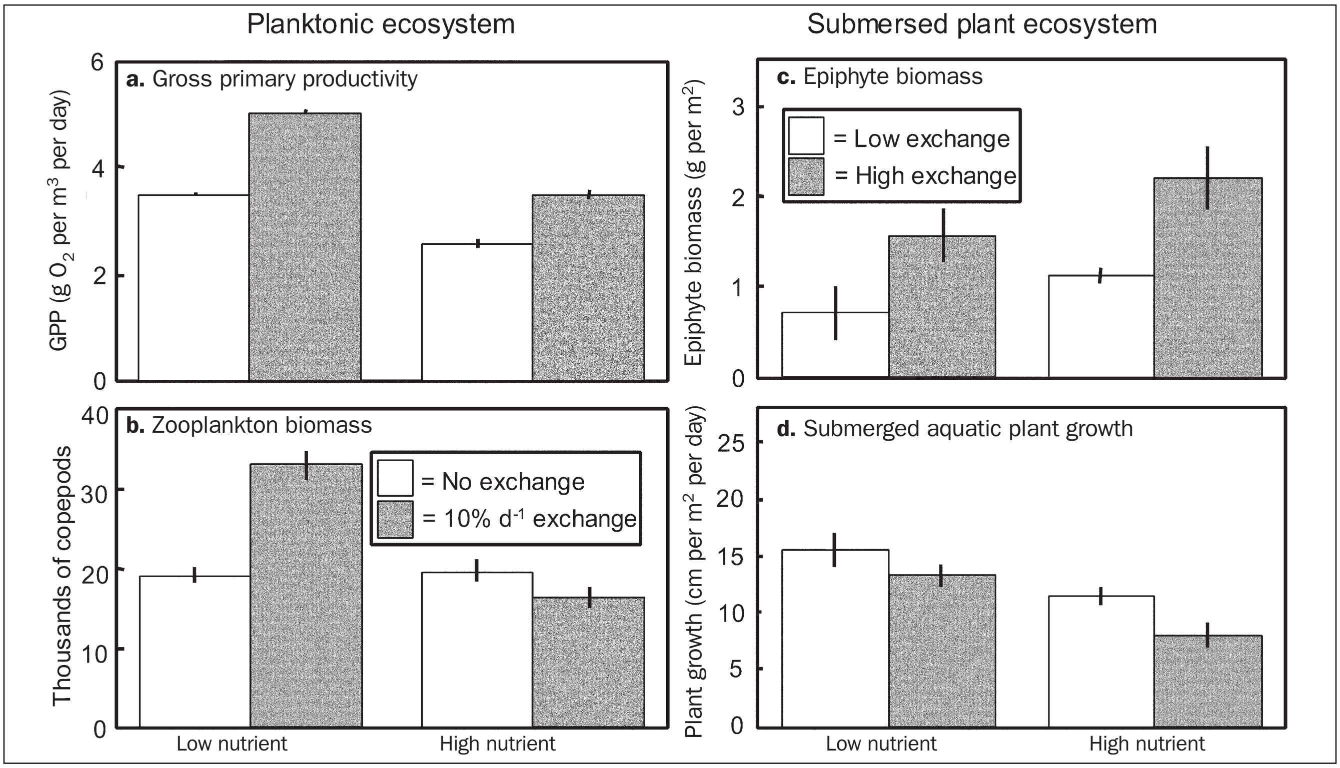


Figure 10. Effects of water exchange rate on aquatic producers and consumers. (a) Effects of 0% and 10% per day exchange of high and low nutrient exchange water on phytoplankton gross primary productivity (GPP, in grams of molecular oxygen per cubic meter per day). (b) Effects of the same treatements on zooplankton abundance (thousands of copepods per cubic meter). (c) Effects of 100% and 1200% per day exchange of high and low nutrient exchange water on epiphyte biomass (grams per square meter). (d) Effects of the same treatment on submerged aquatic plant growth (centimeters per square meter per day). In all cases, bar height represents daily measures averaged over 30-day experiments. Error bars are standard error of the mean among replicates.

series of tanks, each of which represents a different trophic level, a different habitat, or a different distance from the mouth of an estuary (e.g., Adey and Loveland 1991). To examine how processes within submerged aquatic vegetation (SAV) beds interact with those in the deeper water column, we compared the behavior of coupled mesocosms with different types of exchange between SAV and planktonic–benthic habitats (figure 4d; e.g., Murray et al. 2000). The 1-m^3 planktonic–benthic habitat was identical in dimensions and mixing to the C tanks described previously (figure 5). The SAV habitat was a 1-m^3 rectangular tank (1.50 m long, 1.00 m wide, 0.67 m deep), mixed with a paddle wheel. An Archimedes screw pump was used to control the rate of bidirectional water exchange between the planktonic–benthic and SAV components without damaging plankton (visible in the center of figure 4d).

In 45-day experiments, nutrient concentration, epiphyte accumulation, and phytoplankton biomass all increased with exchange rate. We found that, at similar nutrient concentrations, the effects of exchange on this coupled SAV community were more pronounced than the effects on SAV in similar mesocosms (described above) receiving a one-way exchange of filtered estuarine water. Evidently, experiments with one-way flow fail to capture certain dynamics that result from bidirectional interactions. In general, the results of these exchange experiments indicate that the time scale and quality of exchange are critical determinants of experimental outcome. Exchange must therefore be carefully considered in the design of experimental ecosystems and in the comparison and extrapolation of results among systems that differ in exchange.

Effects of physical scale on within-system and among-system variability

Thus far we have emphasized the effects of various aspects of scale on average ecological properties. Scale can also affect variability observed within ecological systems (within-system variability) and variability observed among parallel or replicate ecosystems (among-system variability). Within-system variability occurs over both space (i.e., internal patchiness) and time (fluctuations in dynamics over time). In field studies, as plot size increases, researchers capture an increasing diversity of organisms and habitats; within-system variability increases with increasing spatial extent until all habitat types within the landscape are encompassed.

Applying this to mesocosm research, if our experimental ecosystem is small relative to patch size in nature, then we may only be able to realistically incorporate one habitat type within the mesocosm. If diversity begets stability (a question still subject to much debate; Symstad et al. 2003), then increases in internal heterogeneity associated with increases in size should lead to decreasing internal variability with increases in size; variability over time and among replicates should decrease with increasing spatial extent of the experiment. This logic suggests two related hypotheses that were examined in MEERC mesocosm studies. First, for among-system variability, we postulated that as mesocosm size increases, stability should also increase, leading to decreases in variability among replicates. Second, if we collected multiple samples at different points within a small mesocosm, we would expect to find less variability among these samples than if we followed the same procedure in a larger system; within-system variability should increase with size.

We tested our hypothesis for scale effects on among-system variability using data from planktonic–benthic mesocosm experiments. To illustrate the results, we plotted the coefficient of variation among three replicate mesocosms of each dimension for a combination of measured phytoplankton-related properties (chlorophyll *a*, phaeophyton, particulate organic carbon and nitrogen, and dissolved silica) versus a characteristic mesocosm length scale (we used the square root of total mesocosm surface area). We found that the coefficient of variation decreased exponentially with increases in mesocosm length (figure 11a). From the standpoint of experimental design, this means that, on average, fewer replicates should be needed in large than in small mesocosms to achieve the same degree of statistical power. Although we attribute the pattern to logistic constraints rather than to statistical consideration, it is interesting to note that replication does, indeed, decrease with increasing mesocosm and plot size (e.g., figure 2a; see also Kareiva and Andersen 1988).

To test our second hypothesis for scale effects on within-system variability, we sampled copepod abundance at five evenly spaced points across the diameter of mesocosms of three sizes in the anchovy-grazing experiments described above. We found that the within-system coefficient of variation for zooplankton abundance did indeed increase with container length scale (figure 11b). This finding might be attributed to increasing variability of mixing within larger mesocosms or to patchiness associated with animal behavior. Further, this tendency for within-system variability of zooplankton to increase with container length scale appears to be accentuated as experimental duration increases (figure 11b). These results suggest that the scale dependence of variance for key ecosystem properties follows general patterns and that consideration of these patterns may be useful for improving the design and interpretation of ecological experiments.

Designing scale-sensitive experiments: Rules and tools

The multiscale experiments described in this article have been explicitly designed to examine how ecosystem dynamics vary with scale. Our results all point to the conclusion that scaling choices can profoundly affect experimental outcome. We recognize that, unlike those described here, most experiments are conducted at a single scale and are designed to address questions other than the effects of scale. Recognizing that scale is important, how might ecological researchers apply the best available theory and empirical knowledge to design "scale-sensitive" *(sensu* Bissonette 1997) experiments?

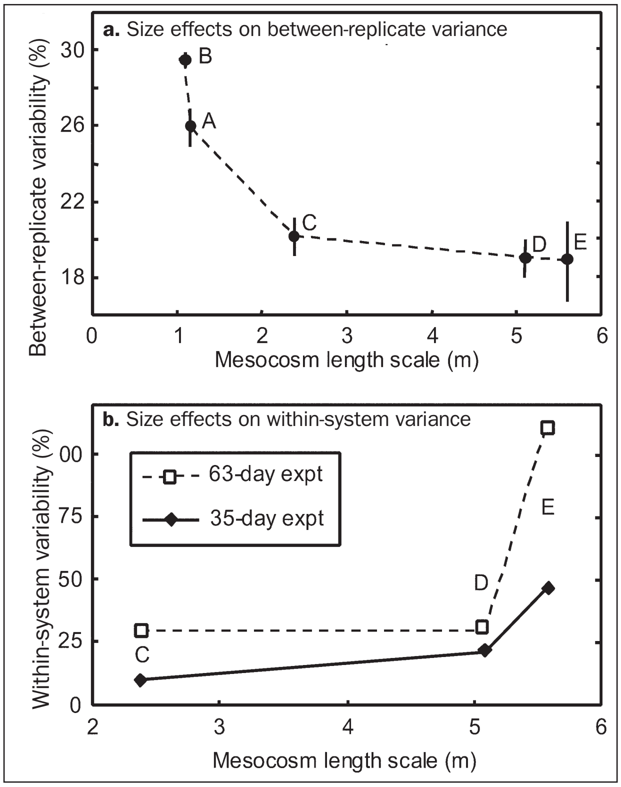


Figure 11. Between- and within-replicate variation as a function of mesocosm length scale. The coefficient of variation was calculated as $100 \times \sum[(X - X_{bar})/X_{bar}]$. *(a) Effects of scale on between-replicate variation. Data are an average of phytoplankton-related properties. (b) Effects of scale on within-replicate variation. Data are for variability in zooplankton abundance within each dimension tank. The* x*-axis is a derived length scale (in meters) calculated as the square root of the total horizontal and vertical surface areas of each mesocosm.*

A first principle in designing scale-sensitive studies is to maximize consistency among three scales: (1) scales associated with the research goals, (2) scales associated with the organisms and processes under investigation, and (3) scales associated with the physical conditions of the experiment. This means explicitly designing the experiment to create a match between experimental and characteristic scales. For instance, for many studies, it is appropriate to choose the duration of the study (i.e., experimental extent) as a whole-integer multiple (at least one or two) of the generation times of the dominant organisms (i.e., a characteristic temporal extent). This practice ensures that the change observed during the experiment is not biased by starting and ending at different phases in the organisms' life cycle. Likewise, it is important to consider the home range (i.e., the characteristic spatial extent) of the dominant organism in selecting mesocosm size (i.e., the experimental extent). Experimental designs that fail to match the experimental scales with the characteristic scales of key organisms and processes frequently result in erroneous conclusions (Tilman 1989).

Dimensional approaches to scaling experimental ecosystems. We have asserted that mesocosms are inherently distorted representations of nature. A key question, then, is whether researchers can somehow compensate for these distortions in the design and interpretation of experiments. The term *dimensional analysis* encompasses a variety of techniques that are based on the proposition that certain universal relationships should apply regardless of the dimensions of a particular system under investigation. In general, these techniques involve developing dimensionless relationships that capture the balance between processes or forces governing the dynamics of a particular system. Dimensional analysis has been widely used for some time in engineering to develop scale models. A number of ecologists have advocated its use as a tool for systematically adjusting for scaling distortions in the design and interpretation of mesocosm studies (e.g., Kemp et al. 1980, Petersen and Hastings 2001). In this case, the central idea is that distortions in one dimension or variable can be counterbalanced by distortions in others to conserve desirable relationships. Below, we offer two examples of how dimensional analysis can be applied in the design of experiments and one example of how this approach can be considered in the interpretation of experiments.

Retaining tidal extent by manipulating slope and porosity in marsh mesocosms. In nature, tidal marshes may extend for miles between upland areas and open coastal water. Enclosed experimental marshes are obviously much more constrained in their spatial extent. How might one design an estuarine marsh mesocosm of modest length (1 to 10 m) that retains both the groundwater residence time and the tidal amplitude of a much longer natural marsh? Darcy's law governs flow through the marsh substrate ($Q = k_p(\Delta H/L)$, where Q = flow rate per unit cross-sectional area, k_p = hydraulic conductivity of the marsh sediment, ΔH = rise in height between the upland border and open-water edge of the marsh, and L = length of marsh. Dimensional analysis suggests that one approach to maintaining similar residence time in a miniaturized marsh mesocosm is to decrease the hydraulic conductivity of the marsh, perhaps by altering the porosity of the sediment media or by introducing other impediments to flow, such as semipermeable membranes. It is possible that modification of this variable generates unacceptable distortions in sediment biogeochemistry, but at a minimum, the dimensional approach makes the scaling tradeoffs involved in design decisions explicit and quantitative rather than implicit and qualitative.

"Fish dipping" as a time-for-space substitution. The "mean-field approximation" is a common implicit assumption that spatial or temporal variability is not important and that the effects of variable conditions can be adequately simulated using average conditions. For example, although light intensity in natural ecosystems is variable on a number of time scales, in mesocosms researchers often employ artificial light of fixed intensity using a fixed light–dark cycle (e.g., 12:12). Likewise, continuous nutrient additions are often used to simulate nutrient input that is variably delivered in nature.

There is increasing recognition that temporal and spatial variability are important in driving ecological dynamics, and that ecological processes that exhibit nonlinear responses, such as the interaction between photosynthesis and light or between predation and prey abundance, do not respond realistically to mean-field approximations. A fundamental problem is that key scales at which variability occurs in nature are often larger than the extent scales of experimental ecosystems. For instance, in the natural pelagic environment, patches of zooplankton are periodically consumed by schools of planktivorous fish that leave a wake of copepod carnage and patches of high-nutrient excreta behind them. In contrast, for relatively small and confined experimental systems, continuous grazing by planktivorous fish that are stocked at mean estuarine densities is likely to result in ecological dynamics that differ from those in nature.

One option for increasing the realism of predator–prey dynamics would be to simulate the patchiness of schooling fish by periodically adding and removing groups of fish from a mesocosm that contains zooplankton (e.g., Harass and Taub 1985). We explored the efficacy of this approach in a series of experiments. In parallel treatments, a small number of anchovies (4 per m^3) were maintained continuously in one set of mesocosms, while a larger number (14 per m^3) were periodically added and removed (using lift nets) from a second set of mesocosms. The frequency and duration of pulsed fish addition was adjusted so that the total number of "fish hours" (fish density x time in tank) was identical in both treatments (approximately 3000 fish hours per m^3). Fish used for the pulse treatment were maintained in holding tanks between additions. We found that copepod abundance was significantly lower in the tanks in which fish were continuously present relative to those that received pulsed addition of predators. Apparently zooplankton populations were able to recover from the intense but short losses to predatory pulses. From this experiment, we conclude that the mean-field approximation can indeed distort experimental results, but also that dimensional manipulations, such as the time-for-space substitution described here, can be effectively used to compensate for such distortions.

Nondimensionalizing time to explore transient responses. This article began with a discussion of the necessary balance between control, realism, and generality in experimentation. In addition to aiding in experimental design, dimensional analysis might be used to increase the generality of experiments by allowing researchers to compare results among experiments that differ in time, space, and complexity. For instance, many ecosystems, both experimental and natural, exhibit transient blooms in primary productivity following initiation or during early succession. These blooms occur on time scales that correspond with the characteristic scales of the organisms involved; phytoplankton reproduce on hour to day time scales, whereas trees reproduce on decade to century time scales. Researchers interested in exploring general theories often choose to use microbial communities with small characteristic time and space scales. Nondimensionalization of time scales may provide a quantitative means of considering the implications of such experiments for ecosystems dominated by larger organisms with slower generation times. For example, although the absolute time scales are quite different, similar bloom patterns are evident in our marsh experiments and our plankton experiments when time is expressed in units of producer generation time (figure 12). For both sets of experiments, the pattern most likely results from depletion of initial resource (nutrient) pools. We found that in plankton-dominated mesocosms this "boom–bust" bloom cycle was substantially damped with continual addition of nutrients.

Simulation models as aids in design and extrapolation. As the examples we have included in our discussion of MEERC experiments suggest, we believe that simulation models provide a valuable tool for addressing questions of scale in mesocosm studies. Both experimental ecosystems and simulation models are abstractions and simplifications of nature, but they represent different balances among control, realism, and generality. To some extent, the advantages and disadvantages of the two approaches are counteracting and

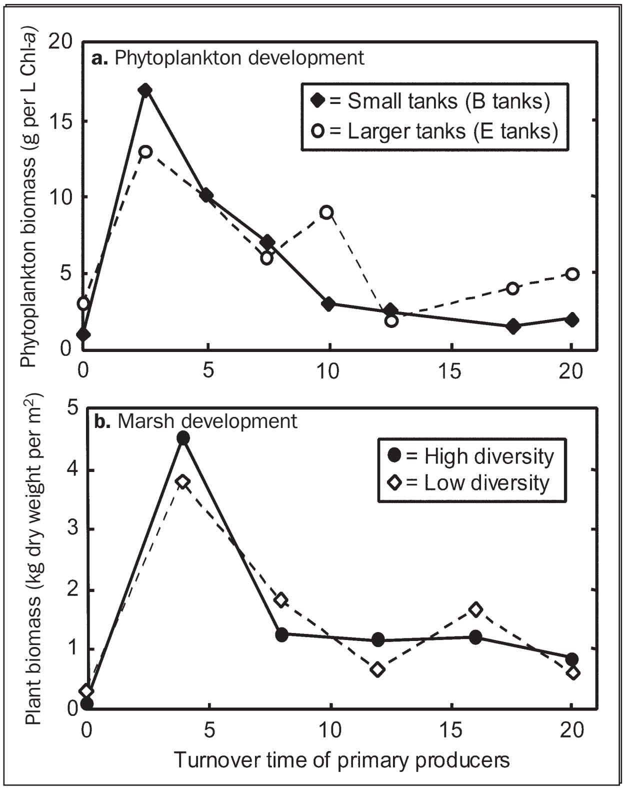


Figure 12. Time series of biomass in (a) phytoplankton (grams per liter of chlorophyll a*) and (b) marsh experiments (kilograms dry weight per square meter), graphed with the same nondimensional time units. Time units for the* x*-axis = (time elapsed in experiment)/(characteristic turnover time for primary producers). Turnover times used were 1.4 days for phytoplankton and 3 months for marsh plants.*

therefore complementary (Kemp et al. 1980). This is evident in a number of instances in which mesocosm studies and simulation models have been employed in parallel (e.g., Parsons 1990). We have found simulation models of great value in interpreting and extrapolating results of the multiscale experiments described in this article. For instance, we developed a model to simulate interactions between plankton, benthic, and periphyton communities and calibrated it using data from our size–shape experiments (figure 5; Petersen et al. 1997, Chen 1998). Simulations revealed that under our experimental conditions, artifacts associated with periphyton growth on mesocosm walls could be reduced to acceptable levels by cleaning the walls twice each week and using tanks with a radius of 2 m or greater. Once the model was calibrated with experimental data, simulations were then run using no-wall scenarios to remove the effect of this artifact and to explore possible effects in nature more realistically (Chen 1998).

Given the importance of spatial heterogeneity in controlling ecological dynamics, we believe that coupling mesocosms with spatially explicitly dynamic simulation models will become an increasingly powerful approach to ecological research (figure 13). In this approach, mesocosms can be thought of as individual cells (grain) within a heterogeneous matrix of different habitats that cover broad spatial extent. Likewise, models can be used to explore effects of temporal variability that are difficult to incorporate in the design of mesocosm studies. We argued earlier that scale effects are often nonlinear and may sometimes exhibit threshold effects. Numerical models offer an excellent tool for exploring nonlinear feedback effects at scales that are larger than individual mesocosms.

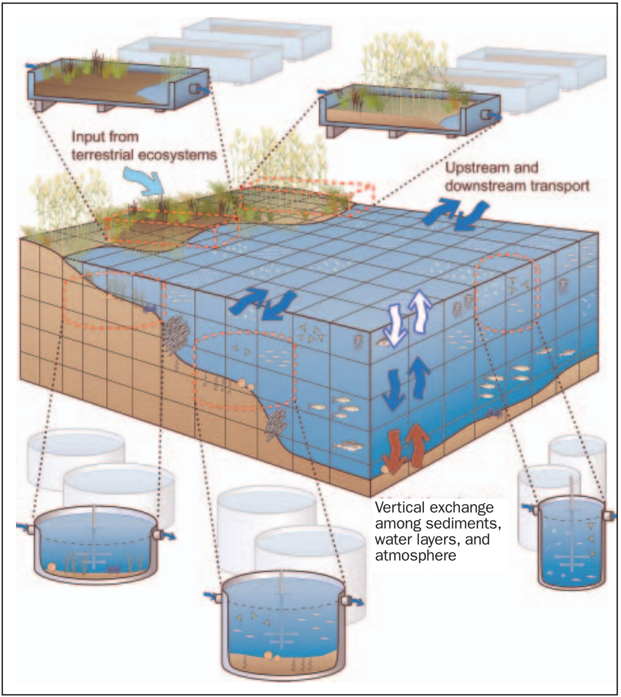


Figure 13. Estuarine ecosystems are heterogeneous in terms of habitat, process, and organism diversity and abundance. Spatially explicit simulation models are a valuable tool for integrating research conducted in the field and in experimental ecosystems in order to develop a more complete understanding of the effects of spatial pattern on process. Water advects, and organisms selectively migrate, among adjacent cells within the aquatic landscape. If one or more grid cells within this landscape are isolated from surrounding cells by walls that limit exchange, the result becomes a conceptual, mathematical, or physical model of an experimental ecosystem.

Conclusion

Mesocosms have proved to be valuable tools for ecological research and are likely to continue to gain in popularity. At the same time, the "science of scale" is increasingly recognized as an important field of ecological inquiry. These two trends should not occur in isolation. Ten years of research at MEERC have revealed that mesocosm experiments can be designed to test a wide variety of scale-related hypotheses that would be difficult or impossible to test in nature. In general, we have found that multiscale experiments can be used both to elucidate fundamental effects of scale, which are evident in both natural and experimental ecosystems, and to quantify artifacts of enclosure, which are solely attributable to the artificial environment in mesocosms. Specifically, we have found that

- Simple geometric and ecological considerations are a good basis for developing scaling hypotheses; however, empirical study can reveal unanticipated effects.
- Direct extrapolation of scaling relationships elucidated in mesocosms to nature is sometimes possible, but experimental findings should be verified at large scales using field data.
- For a given question and system, an "experimental window" of scale, in which mesocosms exhibit realistic ecological dynamics, can often be empirically determined.
- Researchers must be particularly cautious when attempting to extrapolate findings from simple, generalized experimental ecosystems to specific natural ecosystems.
- Rate, quality, and variability in external exchange are critical components of scale that affect ecological dynamics.
- Both within-system and among-system variability in ecological properties are functions of experimental scale.
- Dimensional analysis and numerical simulation models provide valuable tools for designing scale-sensitive mesocosms and extrapolating results.

Other researchers are finding that experimental scale affects diverse ecological relationships, ranging from the ecological impact of toxins (Morris et al. 1994) to vertical transport (Sarnelle 1999) to food chain interactions (Bergström and Englund 2002). Multiscale experiments such as these represent a logical approach to identifying mechanisms by which patterns and processes vary with scale. On the one hand, understanding these mechanisms is crucial to improving researchers' capacity to systematically extrapolate findings from mesocosm studies to whole ecosystems in nature. On the other hand, these experiments also hold the promise of advancing environmental scientists' ability to interpolate information among natural ecosystems that differ in size, shape, and ecological complexity.

Acknowledgments

Our research was supported by the Environmental Protection Agency's STAR (Science to Achieve Results) program as part of the Multiscale Experimental Ecosystem Research Center (MEERC) at the University of Maryland's Center for Environmental Science. Numerous colleagues, students, and technicians contributed to this large research effort, including Joel Baker, Mine Berg, Bob Costanza, Bill Dennison, Pat Glibert, Tim Goertemiller, Todd Kana, Victor Kennedy, Marlon Lewis, Chris Madden, Jeff Merrill, Kelly Merrill, Jenny Purcell, Enrique Reyes, Alison Sanford, Diane Stoecker, Brian Sturgis, Frieda Taub, and Tom Wazniak. The MEERC program built on lessons learned in other mesocosm facilities, most notably the substantial body of work conducted at the Marine Ecosystem Research Laboratory at the University of Rhode Island. Our program benefited from the stimulation provided by a science advisory committee composed of many of the leading thinkers on issues of scale and experimentation.

References cited

Adey WH, Loveland K. 1991. Dynamic aquaria: Building living ecosystems. San Diego: Academic Press.

Bergström U, Englund G. 2002. Estimating predation rates in experimental systems: Scale-dependent effects of aggregative behaviour. Oikos 97: 251–259.

Bissonette JA. 1997. Scale-sensitive ecological properties: Historical context, current meaning. Pages 4–31 in Bissonette JA, ed. Wildlife and Landscape Ecology: Effects of Pattern and Scale. New York: Springer.

Brownlee DC, Jacobs F. 1987. Mesozooplankton and microzooplankton in the Chesapeake Bay. Pages 217–269 in Majumdar SK, Hall LWJ, Austin HM, eds. Contaminant Problems and Management of Living Chesapeake Bay Resources. Philadelphia: Pennsylvania Academy of Science.

Carpenter SR, Kitchell JF, Hodgson JR. 1985. Cascading trophic interactions and lake productivity. BioScience 35: 634–639.

Chen C-C. 1998. Wall effects in estuarine mesocosms: Scaling experiments and simulation model. PhD dissertation. University of Maryland, College Park.

Chen C-C, Petersen JE, Kemp WM. 1997. Spatial and temporal scaling of periphyton growth on walls of estuarine mesocosms. Marine Ecology Progress Series 155: 1–15.

———. 2000. Nutrient uptake in experimental estuarine ecosystems: Scaling and partitioning rates. Marine Ecology Progress Series 200: 103–116.

Crawford SM, Sanford LP. 2001. Boundary shear velocities and fluxes in the MEERC experimental ecosystems. Marine Ecology Progress Series 210: 1–12.

Fee EJ. 1979. A relation between lake morphometry and primary productivity and its use in interpreting whole-lake eutrophication experiments. Limnology and Oceanography 24: 401–416.

Fee EJ, Hecky RE. 1992. Introduction to the Northwest Ontario Lake Size Series (NOLSS). Canadian Journal of Fisheries and Aquatic Sciences 49: 2434–2444.

Frost TM, DeAngelis SM, Bartell DJ, Hall DJ, Hurlbert SH. 1988. Scale in the design and interpretation of aquatic community research. Pages 229–258 in Carpenter SR, ed. Complex Interactions in Lake Communities. New York: Springer-Verlag.

Harass MC, Taub FB. 1985. Effects of small fish predation on microcosm community bioassays. Pages 117–133 in Cardwell RD, Purdy R, Bahner RC, eds. Aquatic Toxicology and Hazard Assessment. Philadelphia: American Society for Testing Materials.

Hartman KJ, Brandt SB. 1995. Predatory demand and impact of striped bass, bluefish, and weakfish in the Chesapeake Bay: Application of bioenergetics models. Canadian Journal of Fisheries and Aquatic Sciences 52: 1667–1687.

Heath MR, Houde ED. 2001. Evaluating and modeling foraging performance of planktivorous and piscivorous fish: Effects of containment and issues of scale. Pages 191–244 in Gardner RH, Kemp WM, Kennedy V, Petersen JE, eds. Scaling Relations in Experimental Ecology. New York: Columbia University Press.

Ives AR, Foufopoulos J, Klopfer ED, Klug JL, Palmer TM. 1996. Bottle or big-scale studies: How do we do ecology? Ecology 77: 681–685.

Kareiva P, Andersen M. 1988. Spatial aspects of species interactions: The wedding of models and experiments. Pages 38–54 in Hastings A, ed. Community Ecology. New York: Springer-Verlag.

Kemp WM, Lewis MR, Cunningham FF, Stevenson JC, Boynton WR. 1980. Microcosms, macrophytes, and hierarchies: Environmental research in the Chesapeake Bay. Pages 911–936 in Giesy JP, ed. Microcosms in Ecological Research. Springfield (VA): National Technical Information Service.

Kemp WM, Petersen JE, Gardner RH. 2001. Scale-dependence and the problem of extrapolation: Implications for experimental and natural coastal ecosystems. Pages 3–57 in Gardner RH, Kemp WM, Kennedy V, Petersen JE, eds. Scaling Relations in Experimental Ecology. New York: Columbia University Press.

Levin S. 1992. The problem of pattern and scale in ecology. Ecology 73: 1943–1967.

Meentemeyer V, Box EO. 1987. Scale effects in landscape studies. Pages 15–34 in Turner MG, ed. Landscape Heterogeneity and Disturbance. New York: Springer-Verlag.

Merrell JR. 1996. Effects of spatial scaling and predation by copepods on microzooplankton population dynamics. Master's thesis. University of Maryland, College Park.

Morris RG, Kennedy JH, Johnson PC, Hambleton FE. 1994. Pyrethroid insecticide effects on bluegill sunfish in microcosms and mesocosms and bluegill impact on microcosm fauna. Pages 373–395 in Graney RL, Kennedy JH, Rodgers JH Jr, eds. Aquatic Mesocosm Studies in Ecological Risk Assessment. Boca Raton (FL): CRC Press.

Mowitt WP. 1999. Scale-dependence of bay anchovy *(Anchoa mitchilli)* growth and top-down control by anchovies of plankton communities in estuarine mesocosms. Master's thesis. University of Maryland, College Park.

Murray L, Sturgis RB, Bartleson R, Severn W, Kemp WM. 2000. Scaling submersed plant community responses to experimental nutrient enrichment. Pages 241–257 in Bortone S, ed. Seagrasses: Monitoring, Ecology, Physiology, and Management. New York: CRC Press.

Naeem S, Li SB. 1998. Consumer species richness and autotrophic biomass. Ecology 79: 2603–2615.

Newberger TA, Houde ED. 1995. Population biology of bay anchovy *Anchoa mitchilli* in the mid Chesapeake Bay. Marine Ecology Progress Series 116: 25–37.

Odum EP. 1961. The role of tidal marshes in estuarine production. Conservationist 15: 12–15.

———. 1984. The mesocosm. BioScience 34: 558–562.

Parsons TR. 1990. The use of mathematical models in conjunction with mesocosm ecosystem research. Pages 197–210 in Lalli CM, ed. Enclosed Experimental Marine Ecosystems: A Review and Recommendations. New York: Springer-Verlag.

Petersen JE, Chen CC. 1999. A method for measuring depth-integrated community metabolism in experimental planktonic–benthic ecosystems. Hydrobiologia 391: 23–32.

Petersen JE, Hastings A. 2001. Dimensional approaches to scaling experimental ecosystems: Designing mousetraps to catch elephants. American Naturalist 157: 324–333.

Petersen JE, Chen C-C, Kemp WM. 1997. Scaling aquatic primary productivity: Experiments under nutrient- and light-limited conditions. Ecology 78: 2326–2338.

Petersen JE, Sanford LP, Kemp WM. 1998. Coastal plankton responses to turbulent mixing in experimental ecosystems. Marine Ecology Progress Series 171: 23–41.

Petersen JE, Cornwell JC, Kemp WM. 1999. Implicit scaling in the design of experimental aquatic ecosystems. Oikos 85: 3–18.

Peterson DL, Parker VT, eds. 1998. Ecological Scale. New York: Columbia University Press.

Porter ET. 1999. Physical and biological scaling of benthic–pelagic coupling in experimental ecosystem studies. PhD dissertation. University of Maryland, College Park.

Porter ET, Sanford LP, Suttles SE. 2000. Gypsum dissolution is not a universal integrator of 'water motion.' Limnology and Oceanography 45: 145–158.

Sanford AM, Morgan JA, Evans D, Ducklow HW. 2001. Bacterioplankton dynamics in estuarine mesocosms: Effects of tank shape and size. Microbial Ecology 41: 45–55.

Sarnelle O. 1999. Zooplankton effects on vertical particulate flux: Testable models and experimental results. Limnology and Oceanography 44: 357–370.

Schindler DW. 1998. Replication versus realism: The need for ecosystem-scale experiments. Ecosystems 1: 323–334.

Schmitz JP. 2000. Meso-scale community organization and response to burning in mesocosms and a field salt marsh. Master's thesis. University of Maryland, College Park.

Schneider DC. 2001. The rise of the concept of scale in ecology. BioScience 51: 545–553.

Short FT, Burdick DM, Kaldy JE III. 1995. Mesocosm experiments quantify the effects of eutrophication on eelgrass, *Zostera marina.* Limnology and Oceanography 40: 740–749.

Smith EM, Kemp WM. 1995. Seasonal and regional variations in plankton community production and respiration for Chesapeake Bay. Marine Ecology Progress Series 116: 217–231.

Staver KW, Brinsfield RB. 1993. Seepage of groundwater nitrate from a riparian agroecosystem into the Wye River Estuary. Estuaries 19: 359–370.

Sturgis RB, Murray L. 1997. Scaling of nutrient inputs to submersed plant communities: Temporal and spatial variations. Marine Ecology Progress Series 152: 89–102.

Symstad AJ, Chapin FS III, Wall DH, Gross KL, Huenneke LF, Mittelbach GG, Peters DPC, Tilman D. 2003. Long-term and large-scale perspectives on the relationship between biodiversity and ecosystem functioning. BioScience 53: 89–98.

Taylor D, Nixon S, Granger S, Buckley B. 1995. Nutrient limitation and the eutrophication of coastal lagoons. Marine Ecology Progress Series 127: 235–244.

Tilman D. 1989. Ecological experimentation: Strengths and conceptual problems. Pages 136–157 in Likens GE, ed. Long-Term Studies in Ecology. New York: Springer-Verlag.

Tilman D, Knops J, Wedin D, Reich P, Ritchie M, Siemann E. 1997. The influence of functional diversity and composition on ecosystem processes. Science 277: 1300–1302.

Wiens JA. 1989. Spatial scaling in ecology. Functional Ecology 3: 385–397.