Chapter 6
Data on Distribution and Abundance: Monitoring for Research and Management

Samuel A. Cushman and Kevin S. McKelvey

In the first chapter of this book we identified the interdependence of method, data and theory as an important influence on the progress of science. The first several chapters focused mostly on progress in theory, in the areas of integrating spatial and temporal complexity into ecological analysis, the emergence of landscape ecology and its transformation into a multi-scale gradient-based science. These chapters weaved in some discussion about the interrelationships between method and these theoretical approaches. In particular, we discussed how powerful computing, large spatial databases and GIS cross-fertilized ecological theory by enabling new kinds of analyses and new scopes of investigation. However, up to this point we have given relatively little attention to the third leg of this triad, data. This and following chapters focus explicitly on data. The next several chapters discuss the advances in broad-scale data collection and analysis enabled by remote sensing, molecular genomics and satellite GPS telemetry, and how these data have made fundamental contributions to virtually all branches of ecology, especially spatial ecology, landscape ecology, and global scale research.

The goal of this chapter is to establish a framework for how data collection and management can best be designed to interact with modeling and analysis across both space and time. The chapter is divided into five sections. First, we discuss the fundamental importance of quality, large sample, spatially referenced, broadly distributed data for reliable inferences to advance research and to guide adaptive management. Second, we explore the challenge posed by limited quality, quantity and extent of data on species and environmental conditions over space and time and discuss the limitations this poses to effective monitoring to guide research and adaptive management. This then provides motivation for a discussion of the importance of monitoring resources themselves, in context and in particulate. We then introduce the concept of the four-dimensional monitoring data-cube, and argue that by collecting accurate data at a fine spatial scale and across large geo-

S.A. Cushman(✉) and K.S. McKelvey
US Forest Service, Rocky Mountain Research Station, 800 E Beckwith, Missoula, MT 59801, USA
e-mail: scushman@fs.fed.us

graphical extents for multiple resources we will be able to produce a flexible, multi-variate, multi-scale data structure that will optimally support ecological analysis and meet monitoring needs for adaptive management. Finally, we discuss linking gradient modeling and integrated, multiple-scale monitoring. Because ecological systems are highly complex and vary dramatically across space and time, we need to think differently about data collection, monitoring and statistical analysis. As we discussed in Chapter 1, monitoring and analysis should not attempt to obtain replicated samples of representative individuals from unstructured populations, because such unstructured populations do not exist and independent sampling in space and time is usually impossible. Rather, the goal is to directly integrate space and time into a sampling frame, as described in the four-dimensional data cube idea described below, and link this to flexible, gradient modeling to infer condition and trend of ecological attributes over space and time.

6.1   Monitoring for Research and Adaptive Management

Adaptive management works by specifying resource goals, conducting management whose purpose is to create or maintain these desired conditions, and monitoring results to confirm that the system is behaving as expected and that resources are moving toward the desired conditions. This approach presupposes that the state of the system is well known across time. A good example of adaptivemanagement occurs whenever you drive a car. The general direction of travel is relatively constant, but, based on visual data, you constantly make small adjustments to keep the car on the road. Because you have precise data concerning where the road lies relative to your current direction of travel, this is easy and effective. However, if you were driving in dense fog or were driving blindfolded, you would quickly crash. For tracking the trajectories of ecological systems, monitoring data serve the same purpose as your eyes while driving. As a result, monitoring resource condition and trend has greatly elevated importance under the adaptive management paradigm. Cost effective, timely, representative, and broad-scale monitoring of multiple resources is the foundation on which adaptive management depends. Adaptive management literally cannot be "adaptive" without these data.

The adaptive management paradigm sets high priority on developing ongoing analyses, based on monitoring, to continually adjust or change land management planning decisions and thereby efficiently move toward desired conditions. The adaptive management cycle involves: (1) a comprehensive evaluation of current resource conditions, (2) frequent monitoring and evaluation of condition and trend relative to desired conditions, and (3) adaptation of management to improve performance in approaching or maintaining desired conditions. Multiple resource monitoring is critical for establishing ecologically meaningful and appropriate desired conditions, evaluating current conditions relative to these objectives, and evaluating effects of management over time to guide adaptive changes to the management regime. For monitoring to provide meaningful

information to the adaptive management cycle it must provide statistically rigorous measurements of the condition and trend of multiple resources across the analysis area with sufficient temporal frequency to provide the periodic evaluations of resource condition and trend to guide management adaptation.

Broad-scale, large-sample, georeferenced measurement of multiple biological and abiotic environmental attributes is also the foundation for addressing spatial complexity and temporal variability in ecological research, as described in Chapters 2 and 3. In Chapter 3 we discussed the importance of focusing at the scale of organisms and their direct interactions with their environment, and then integrating pattern–process relationships at that grain across large spatial extents. This then implies sampling of species themselves and the environment at multiple scales, including direct measurement of occurrence, abundance, and population dynamics (where possible). The important point here is that, as spatial and temporal complexity are not noise to average away, but fundamentally important attributes of ecological systems, data collection for research and monitoring for adaptive management must be fine-grain, large extent, large sample, georeferrenced, measurement of multiple ecological attributes carefully chosen to directly represent the species, process or attributes of interest. As ecological systems are spatially complex, temporally dynamic, and scale dependent, frequent, multi-scale, spatially referenced data collection is the foundation for understanding. Given the spatial, temporal and contextual nature of ecological systems, frequent remeasurement across large spatial sampling networks is fundamentally important.

There are several critical attributes that a data collection or monitoring program must possess for it to be successful in providing reliable inferences about condition and trend to support analysis of ecological systems in particulate, in context, and integrated over space and time. We suggest that all monitoring initatives and existing programs be evaluated with respect to these essential attributes and that prioritization be given preferentially to monitoring efforts that provide statistically powerful inferences about condition and trend based on representative empirical samples.

Key attributes:

1. Based on empirical samples. For monitoring to provide any reliable information about condition and trend of a resource it must monitor the resource itself, or a proxy that has been reliably shown through rigorous scientific research to be a surrogate for the resource. Given that very little rigorous science exists relating resources to proxies, we strongly favor monitoring of resources themselves.
2. Based on representative samples. Samples must be collected in a representative manner from the target population. Representative sampling is essential to avoid biases in estimates of resource condition and trend. Monitoring inferences based on nonrepresentative samples are of unknown accuracy and thus of limited utility as a guide for adaptive management.
3. Provide sample sizes that are sufficient to provide statistically powerful inferences of condition and trend for the evaluation area at least every 5 years. This is perhaps the most daunting requirement for an acceptable monitoring program. In the past few monitoring efforts have evaluated their statistical

power to describe conditions and detect changes. However, without formal evaluation of statistical power in relation to sample size, analysis area, and temporal sampling period, the information provided by monitoring is of unknown value. Statistical power must be measured a priori to determine if a monitoring effort has the ability of describing current conditions and detecting changes with an acceptable probability. Acceptable confidence intervals and statistical power for estimates of condition will vary among resources given inherent variability of the data, importance and risk.

4. Spatially representative and well distributed. One of the key concepts in this book is that spatial pattern and temporal variation matter fundamentally to ecological processes. Therefore, in monitoring our goal is often not to estimate some mean attribute of some large heterogeneous area, but to measure, map and model the variation in the conditions of multiple resources across broad spatial extents. This has major implications for monitoring. Specifically, it requires that monitoring efforts be spatially informed, with representative sampling stratified by ecological strata across the spatial domain, with large sample sizes obtained without major spatial gaps in distribution, and with sufficient density to reflect the variability of ecological patterns and processes. Often spatial autocorrelation and spatial dependence will be of direct interest (see Chapter 7). In such cases, monitoring and other data collection must be guided by the desired precision in the spatial analyses that will follow, including details about the distances between all pairs of observations, and ensuring that these distances are to some degree optimized to accommodate autocorrelation and semi-variance analyses.

5. Based on recent samples. The age of data is a major issue in monitoring. Ecological conditions, species populations, and management activities all change over time. Data that is many years old has unknown relationships to current resource conditions. Thus, monitoring programs should continually collect new data and should base all inferences on recently collected data (perhaps within 5 years).

6. Include frequent remeasurement. Given the critical role temporal dynamics play in ecological processes (as discussed in Chapter 2), and the foundational role frequent remeasurement plays in adaptive management of ecological systems, it is essential that data collection be frequently repeated, using comparable methods to collect a consistent collection of variables at consistent scales. The frequency must be sufficient to provide meaningful guidance to managers in the adaptive management process, and for ecological research must be frequent enough to provide sufficiently precise tracking of ecological dynamics of both driver and response variables to reliably link mechanisms with responses.

7. Continue for long periods. To guide management and support understanding of ecological systems, it is critically important that data collection and monitoring efforts be maintained and continued over long periods of time, with consistent sampling in space, with comparable methodologies at consistent scales. This implies large and long-term commitments to maintaining data collection programs, with particular prioritization to maintaining permanent networks of spatially referenced sampling plots.

In the driving example, above, it is critical not only to know the current position of the car relative to the road, but also to be able to look ahead. In ecological systems, looking ahead involves modeling. There are 3 broad approaches to modeling the future. The first is simply to project current trends. The second is a statistical approach which uses past situations to infer the likelihood of various futures. The third is process modeling, where the model represents an animated hypothesis concerning the current state of the system and its dynamic properties. Dynamic models are flexible and allow a wide variety of future scenarios to be simulated, but generally lack statistical understandings of their validity. The purpose of this chapter is not to discuss these models in detail, but to note that none of these approaches has any validity without appropriate input data. For example, to project a current trend one must have data of high enough quality to produce the trend in the first place. An additional positive attribute for monitoring data is that it is consistent with both initializing and validating critical models.

6.1.1   Challenge of Limited Amount, Extent, and Quality of Spatially Referenced Ecological Information

The seven criteria listed above pose a major challenge to implementation. There are very few data collection programs which possess most of these attributes. Most data collection in ecological research is narrow in scope spatially and limited to a particular moment in time, or if the goal is to look at change, limited to a few temporal snap-shots. This severely limits ability to integrate pattern-process relationships across large, heterogeneous spatial extents and over time. Similarly, most ecological monitoring to guide natural resources management is severely limited by failure to adequately consider spatial sampling design, often failure to even establish spatially referenced sampling networks, inconsistencies across space and through time on what variables are measured, at what scale, and with what methods.

   The general failure of most past efforts is largely due to issues associated with the nature of ecological data. Ecological processes and the landscapes they create are highly variable across space and time. The community composition, for example, at any location is determined by a myriad of factors at many scales. The microclimate, soils, juxtaposition, specific disturbance history, geographic location, and deeper history all play a part. Areas even a few meters away can be characterized by radically different communities due to the interactions of these factors in different combinations. For example, a slight change in aspect (microclimate) may lead to a site being overrun by an invasive weed if that site has been recently disturbed and if it is proximal to source populations of the weeds, is in an area within the range of the exotic, and if it lies in a continent or island where the native vegetation cannot compete with the weed and if no other weeds have already colonized the area. Most of these factors are both spatially and temporally variable: new road construction provides a proximal source population for the weeds; the weed continues to spread into new regions; a fire, windstorm, or insect outbreak produces necessary levels of disturbance; the local climate changes.

The highly variable and intrinsically multivariate nature of these data has historically relegated broad-scale monitoring to the collection of coarse data at large spatial scales. For example, in the weed example above, a national grid of vegetation plots with one plot every 10 km would only be able to directly speak to the weed spread at a broad scale. For example, in the county in which we reside, there would only be about 67 plots. If the climatic zone in which the weed could invade represented 10% of the total land area in the county, spread statistics would be based on at most 6–7 plots. Implementing this coarse grid, however, would be neither easy nor cheap. Nationwide there would be over 91,000 plots. At US$1,000 per plot (which is less than current vegetation monitoring systems cost per plot), this operation would require US$90,000,000.00 and a large and diffuse bureaucracy to implement. This combination of coarse data resolution coupled with high cost in turn leads to low levels of support and therefore inconsistent implementation across time and space.

This is a fundamental problem with collecting environmental data, and technologies will not cause it to go away. To address these issues, monitoring has turned to a variety of approaches both to increase the spatial and temporal resolution of the data and to reduce costs. Among these, are approaches based on monitoring system "macro-characteristics" rather than particulate and contextual data on individual species and key abiotic patterns and processes. These efforts often take one of two forms. The first is the surrogate approach to monitoring, in which a small collection of species or attributes are measured in the hope that their dynamics will represent those of the system. The second is the coarse filter approach, in which a few macro-characteristics, such as some broadly defined ecological community types, will provide sufficient information to infer the dynamics of the species and processes that act within them. In the two sections that follow we argue that neither of these approaches typically is sufficient to reliably track or predict the ecological condition and dynamics of populations and ecological processes across space and through time.

6.2 Species Surrogate Approaches

Species surrogacy has a long history in the field of conservation biology (Landres et al. 1988; Lambeck 1997; Wiens et al. 2008), with a number of different variants and implementations of the concept. Variants include ideas associated with shared habitat or functional requirements (guild membership), trophic dependencies (keystone species), area requirements (umbrella species), ecological function (engineer species), and ecological associations (focal species as defined by Lambeck 1997; See Noon et al. 2008 for a more complete list of surrogate types). Recently, Wiens et al. (2008) proposed that by grouping species using multivariate clustering it is possible to identify surrogates from each of the resulting groups to represent the group for purposes of monitoring. From a historical standpoint this concept is most similar to the guild indicator concept (Block et al. 1987).

The primary driver behind using surrogacy to monitor ecosystems is expediency. As Wiens et al. (2008) state: "…surrogate species or groups of species can be used as proxies for broader sets of species when the number of species of concern is too great to allow each to be considered individually." Implicit in this, of course, is the assumption that the surrogacy approach is effective. There have been few real tests of this, but when done, the results are seldom encouraging (Verner 1984; Landres et al. 1988; Andelman and Fagan 2000; Lindenmayer et al. 2002; Roberge and Angelstam 2004). This is obviously less onerous than monitoring all species individually, but entails large risks of bias if the chosen indicator does not sufficiently represent the abundances of other species within the group.

Cushman et al. (in press) evaluated the surrogate species concept for forest birds at two spatial scales and under two species grouping approaches. The overarching question was whether the abundance of a species across a large sample of locations provides a surrogate for the abundance of other species. They evaluate this based on an a priori grouping of species into life-history categories (Hansen and Urban 1992), and on an empirical grouping based on observed similarity of abundance patterns. In the first case, they tested whether there are strong surrogate relationships among species within objectively defined groups formed on the basis of ecological characteristics. In the latter case, the test was whether species in empirically formed groups derived from cluster analysis provide substantial surrogacy for the abundances of other group members. Data consisted of 72,495 bird observations on 55 species across 1,046 plots distributed across 30 sub-basins. They analyzed abundance patterns at two spatial scales (plot and sub-basin) and for two grouping rules. There were few significant indicator relationships at either scale or under either grouping rule, and those few found were unable to explain a substantial portion of the abundance of other species. They concluded that, coupled with the lack of proven efficacy for species surrogacy in the literature, these results indicate that the utility of indicators and similar types of surrogate approaches must be demonstrated rather than assumed.

Surrogacy has served as the lynchpin of ecosystem management and the movement away from single species paradigms. However, effective species surrogates appear to be rare. Fundamental ecological theory offers a possible explanation. No two species can long occupy the same niche (Gause 1934; Hutchinson 1957; Pulliam 2000). Thus, all coexisting, sympatric species must differ along at least one critical niche dimension. There must be some limit to the similarity of coexisting species (MacArthur 1967), and it is expected that species that are similar in some aspects of their niche will displace on others so to minimize competition. This would tend to lead to weak or negative patterns of co-occurrence for species sharing functional ecological characteristics, as was found in Cushman et al. (in press). Niche displacement processes would appear destructive to the stable existence of strong species surrogacy based on functional ecological characteristics.

As discussed in Chapter 2, community patterns are usually nonequilibrial in space and time, with frequent reversals in competitive balance, changes in relative fitness, the severity of predation and parasitism through time and over complex

landscapes. Given this inherent disequilibrium in time and variability through space, simplistic ideas of species surrogacy seem unlikely to hold. The few evaluations, such as Cushman et al. (in press), have typically found poor performance of species surrogates. We believe that given the individualistic nature of species responses to scale dependent environmental and biotic interactions, it is necessary to consider ecological systems in particulate and in context, rather than seek surrogates to explain large components of the unsampled system.

6.2.1   Community Type as Surrogate for Population and Process

Coarse filter approaches to conservation assessment are attractive because of their potential efficiency and low cost. They attempt to place many species under the umbrella of one habitat assessment effort, and efficiency is obtained by monitoring a mosaic of community types as a surrogate for species viability. However, it is widely recognized that the sufficiency of coarse filter approaches are largely untested.

   For community types to succeed as a coarse filter proxy for population performance at least four conditions have to be met simultaneously.

1. Habitat must be a proxy for population performance.
2. Mapped community types must be a proxy for habitat. Suitable habitat for any given species will involve ranges of critical resources and conditions across several spatial scales (Cushman and McGarigal 2004). Coarse filter conservation approaches are based on managing landscapes to provide certain amounts or configurations of vegetative community types, a great simplification both of habitat and of scale. Thus, for the coarse filter to proxy for habitat, mapped community types must be strongly related to species distributions, abundance, and population performance. Specifically, they must explain a large proportion of the species–habitat relationship for all species involved.
3. Suitable habitat is a species-specific characteristic. Each species, whether aquatic or terrestrial, has unique tolerances for different ranges of environmental conditions and unique requirements for critical resources. For the coarse filter to be successful, the mosaic of discrete community types which define the coarse filter habitat must be a good proxy for the responses of all species of interest.
4. In addition, these relationships must not decouple with management or natural disturbances. Specifically, the proxy of mapped community type for habitat may break down when management or natural disturbance occurs at a scale that will change habitat at a scale that species are directly responding to but not at a scale that results in changes in mapped community types. If this decoupling occurs, assumed correlations between the coarse community types and habitat may become invalid, resulting in the failure of the coarse filter to predict species performance. For example, stand-level vegetation manipulation (e.g. fuels treatment) in which the structure alone is modified such as decreasing density of the stand,

can have large effects on many vegetative and abiotic attributes of a stand, but may not change the cover type or seral stage as mapped by the coarse filter. In this case, the coarse filter is insensitive to important changes in habitat, resulting in a decoupling which may impair the sufficiency of the population performance proxy. In this case, additional attributes of forest structure (density or the relative openness of the forest) become important.

The community type coarse filter, therefore, represents four layers of proxies between what is measured and actual population performance. Because all four of these conditions need to be met simultaneously, errors are multiplicative. Consider an example where 75% of the variance in a population's performance is directly attributable to habitat, and, of that, 75% is directly related to broad cover types representing the utilized coarse filter, and coarse/fine scale decoupling only results in a 10% degradation of correlations. These relationships are all optimistic, as few published studies have reported explained variances of this magnitude. Nevertheless, the overall efficacy of the coarse filter will be $75^*75^*90\% = 51\%$ in this example. Thus it is by no means certain that coarse filter approaches will provide the required quality and resolution to perform with sufficient power and resolution to reliably represent the populations and processes they are intended to proxy, even if each level of the proxy chain performs well.

Cushman et al. (2007) conducted a multivariate analysis of bird community relationships to multiple scale habitat data in Oregon and Wyoming to test these assumptions of ecological community type as a proxy for the composition and abundance of the avian community. Their results indicate that the effectiveness of vegetation communities as proxies for population status is tenuous – even with the best, multi-scale vegetation data less than 60% of the variation in species presence is explained – and the effectiveness of landscape composition of vegetation types is highly dependent both on thematic resolution and how "community types" are defined. Their results suggest that, at a minimum, coarse filter community types should be characterized by a combination of detailed cover types across multiple seral stages. In addition, their results show that there are large differences in habitat relationships among species and that a single representation of coarse filter elements is unlikely to be an effective proxy for multi-species habitat. Furthermore, their results indicate that the relationship between habitat coarse filters and species distribution and abundance may decouple following management activities; species distribution and abundance may not be predictable based on mapped habitat in areas that are actively managed if the mapped habitat does not account for diversity in structure in terms of density, size, and multistory vs single story conditions.

## 6.3   Importance of Monitoring Resources Themselves

The arguments and examples given above speak directly to issues associated with using small subsets of species as surrogates, or mosaics of broadly defined cover types as coarse filters for multi-species terrestrial animal population performance.

The issues raised, however, are universal. We have argued that representative data that allow precise and unbiased estimation of ecosystem conditions are essential for adaptive management. We have also noted acquiring these data is both difficult and expensive. Given this, the loss of power and injection of largely unknown biases associated with using surrogate measures will generally lead to unacceptable power degradation. As demonstrated in the example above, this is true even if correlations between the surrogate and the resource of interest are high. In our investigations, and in the vast majority of similar investigations in the literature, these correlations are often weak and not significant. We therefore conclude that effective monitoring of a resource both to estimate its state and track its dynamics will almost always involve direct measurements of that resource. Power loss also occurs when a resource is modeled based on measured values. In a simple example, tree volume cannot be measured directly in the woods. Instead, it is modeled by measuring diameter and height, and by assuming a "form factor" to estimate taper. This is a type of surrogacy in which one attribute, in this case volume, is inferred through the measurement of two other variables. Volume estimates will contain the multiplicative errors associated with both diameter and height measurements and will contain unknown bias due to an unmeasured, but assumed, form factor. When making decisions on what to monitor, the practical ability to produce estimates with the necessary power will be directly linked to the degree to which the desired attributes and conditions can be framed in terms direct, easy, and accurate measurements.

6.3.1   The Multi-variate, Multi-scale Monitoring Data-Cube

In Chapter 5 we discussed a gradient-based paradigm for multiple-scale ecological analysis. The basic idea was that the state of any ecological system will be defined by the interactions of many processes, the behavior and distribution of many organisms, and abiotic environmental patterns across a broad range of spatial scales. There is no practical way to specify the state of the environment as a single attribute (e.g. classified map) that will optimally relate to multiple processes and how they affect the biotic and aiotic components of the system. Instead of seeking a few surrogates or a macro-attribute coarse filter which simultaneously answers all questions, we believe it is more practical and effective to retain multiple variables in the analysis at the native scale of their dominant interaction with the major ecological processes in the system. This is the multi-variate, multi-scale monitoring data cube idea. In this data cube the x, y dimensions represent space, and the z dimension represents multiple biotic entities, process or abiotic variables measured in a large spatial network of georeferenced locations (Fig. 6.1).

Importantly, these biotic entities, process and abiotic variables forming the y-dimension include a broad range of spatially referenced ecological data, from a variety of sources, at a variety of scales. As discussed at length earlier in the chapter, these will usually include measurements of multiple ecological variables across a spatial network of permanent plots. Such data could include vegetation inventory
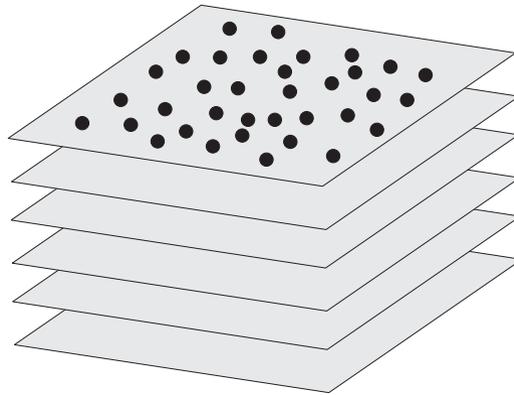
Fig. 6.1 Three-dimensional data cube for flexible multi-scale monitoring. Each layer represents a different ecological sampling dataset, such as point-level vegetation or wildlife occurrence data, topography, climate, soils, remotely sensed data, and disturbance and management history. Each layer is a spatial database, in that all records are represented at their geographical x−y coordinates
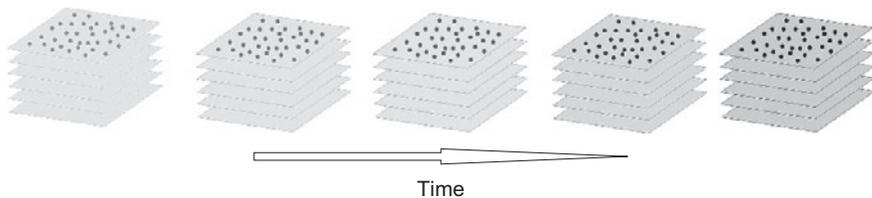
Fig. 6.2 Remeasurement of the three-dimensional data cube produces a four-dimensional data structure. Here we represent the fourth dimension as an arrow of time with remeasured data-cubes stepping out toward the right as time proceeds forward

on permanent vegetation plots, sampling wildlife occurrence or abundance, such as bird point counts, measuring soils depth, texture and chemistry, and recording microclimate variables such as temperature, soil moisture and relative humidity. These y-dimension variables of the data cube also include broader scale information derived from field sampling at stand or watershed scales, as well as spatial data stored as GIS layers on topography, geology, human management history, disturbance history, roads and other development, remotely sensed data such as satellite imagery, lidar, and aerial photography.

The fourth dimension would be repeat measurements of this 3-D data cube over time (Fig. 6.2). This cubic form provides an ability to directly integrate multivariate and spatial analyses, by having multiple measured variables georeferenced across a 2-D geographical space. The linkage to repeat measurements over time provides explicit ability to integrate spatial and temporal processes. The cube is multi-scale, in that the data are re-scaleable to provide optimal multivariate, spatial prediction for a range of dependent variables being inferred or predicted.
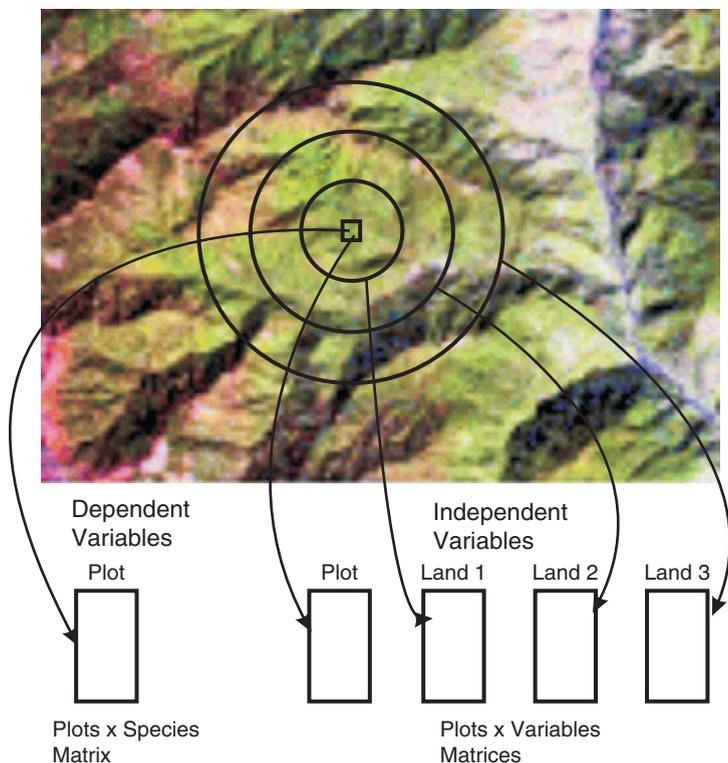
Fig. 6.3 Each layer in the multi-variate, spatial data-cube contains data measured at a particular native grain. Scaling relationships between that layer and a response variable of interest can be calculated by evaluating the strength of the statistical relationship between the response and predictor variable within differing neighborhood sizes and across variation in grain of the data through resampling

This last concept is particularly important, and deserves elaboration. Typically in landscape ecology and natural resources monitoring, what has been done in the past is to collect one or a few environmental attribute layers at a single scale. In landscape ecology it has typically been a single chloropleth classified map of a patch mosaic of landcover types. This data layer usually is the statistical product of a multivariate analysis of several spectral layers or other GIS data. This combination of multiple datasets into a single classified product involves many issues regarding the appropriateness of the classification for different research or monitoring questions, the appropriateness of the scale of the data in terms of grain and minimum mapping unit, the accuracy and meaningfulness of patch boundary definitions for management or research questions. As discussed in Chapter 3 and revisited above, no single-scale classified map can possibly represent the patterns and scales of environmental variability and the shifting importance of multiple organisms and simultaneous processes. This is the prime impetus for the multi-scale data cube (Fig. 6.3). Maintaining the data in a flexible, multi-variate, spatially explicit form facilitates the application of spatial and multivariate statistical models to produce the best predictions (given the
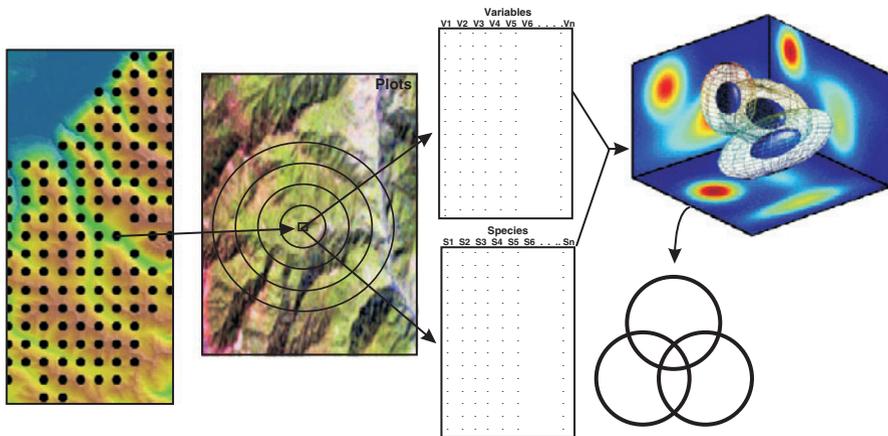
Fig. 6.4 The three-dimensional, spatial data-cube provides an ideal foundation for developing flexible, multi-scale gradient models linking multiple variables at different spatial sales to predictions of species occurrence, ecological process or other ecological attributes

data) for each species, process or attribute without losing information, introducing errors of classification, and collapsing scales (Fig. 6.4).

6.3.1.1   Limitations of Sample-Based Monitoring on Fixed Grids

One of the core concepts behind the multi-variate, spatial data-cube presented above is the central importance of large, spatially referenced, measurement plots for multiple variables. In the past most applications of large permanent re-measure plot systems have focused not on predicting spatial process-pattern relationships but on obtaining non-spatial estimates of mean and variance of some quantitative parameter within some (often large) region of space. It is important to consider the difference between this and the paradigm we are presenting. Instead of trying to estimate the mean value of some quantitative parameter within some spatial unit, our goal would be to adopt a multi-scale gradient perspective (Chapter 5) and represent variability of many ecological attributes simultaneously and continuously across space and through time.

Efforts to estimate mean values for monitored variables within relatively large spatial units are valuable, particularly in the historical sense that they were the impetus for the establishment of the large spatial sampling networks (e.g. FIA) that we presently have. However, they are not sufficient to meet data requirements for a flexible, multi-scale, multi-attribute approach to ecological analysis and adaptive management, as we described in Chapters 2, 3, and 4. First, sample-based statistical trend monitoring usually cannot provide spatial estimates of ecological variables across the analysis area; definitionally they collapse space to increase sample size and therefore power. As discussed in Chapter 3, this kind of upscaling is fraught with potential error and suffers from extreme loss of information, particularly because the spatial pattern of the data are completely ignored. Second, non-spatial sample based estimates of mean values of ecological variables are very often severely limited by sample size to broad,

and often nebulous categories. In traditional FIA based predictions, for example, there is a trade-off between area of inference, sample size, and classification resolution. Thus, for inferences at the National Forest level, in order to obtain a sample size that will provide sufficient statistical power for an acceptable level of precision, it is usually necessary to limit analyses to very broad cover categories, such as cover type or seral stage. Such broad categories have been shown to be poor proxies for the habitat relationships of many wildlife species (Cushman and McGarigal 2004, Cushman et al. 2007). Third, the previous issue of sample size within an analysis area has an inverse dilemma. In order to obtain sufficient power for acceptable precision of a mean estimate for a variable of interest it is also usually necessary to make the analysis area very large. This results in the production of estimates of mean and variance for several ecological variables at very broad scales, scales much broader than usually can be linked to the interactions of ecological entities, such as organisms, and their environment through investigation of processes. Thus, estimating mean values from sample grids simultaneously suffers from two major handicaps, both of which seriously limit ability to analyze ecological pattern-process relationships flexibly across scale.

There are several other issues that limit traditional sampling for trend on fixed grids. First, these efforts are often exceptionally expensive due to the large sample size requirements, which represents a cost barrier to sampling many resources and species. In addition, if detection of population trend is the objective, repeated sampling for many years is usually required before any trend can be detected, even in a common species. Furthermore, sampling for trend provides no explanation for the causes of observed changes, which is essential if we are to use the information to understand ecological processes or guide management. Sampling for trend also provides no means to estimate resource condition across the landscape at areas not directly sampled. For monitoring to be useful for many research and management questions it should provide spatially explicit predictions of conditions across the analysis area. Finally sampling for trend provides no means to predict expected future changes in ecological conditions as a result of changing management or natural disturbance regimes. As discussed above, these issues have collectively served as major disincentives for planning and executing large scale representative monitoring.

## 6.3.2   Gradient Modeling and Integrated, Multiple-scale Monitoring

The above discussion of plot grids has historically presented a conundrum. Huge expense and long time frames are required to obtain data for few variables at scales too large to be useful. Overcoming these limitations, using traditional approaches, would require perhaps an order of magnitude more sampling effort, and would likely still be fraught with many of the same problems. Traditional attempts to overcome these problems through the use of coarse proxies have also proved ineffective. Thus the institutional response has been to do very little monitoring. We therefore find ourselves in the deplorable condition of entering into a period of rapid climate change

with virtually no coherent, spatial baseline data concerning the condition and trends of our major natural ecosystems. This situation, however, could be corrected if there was a way around this historical conundrum. Here we propose a potential response which we believe has merit. The key is to keep all of the collected information, including spatial coordinates, in play (we have visualized this through the four dimensional data cube), and to achieve localization by linking these data through gradient models. Specifically, the combination of this kind of database and gradient modeling can identify the environmental and management factors related to each organism's distribution, and each processes controls, determine the causes of observed changes in ecological conditions, produce spatially explicit maps describing expected conditions across large spatial extents while maintaining grain at the scale of dominant pattern–process relationships, and provide spatially explicit predictions of expected future conditions resulting from altered management and natural disturbance regimes.

In this effort, models developed from current empirical data are used to produce current predictive maps of the condition of particular ecological variables continuously across the analysis area (Fig. 6.5).
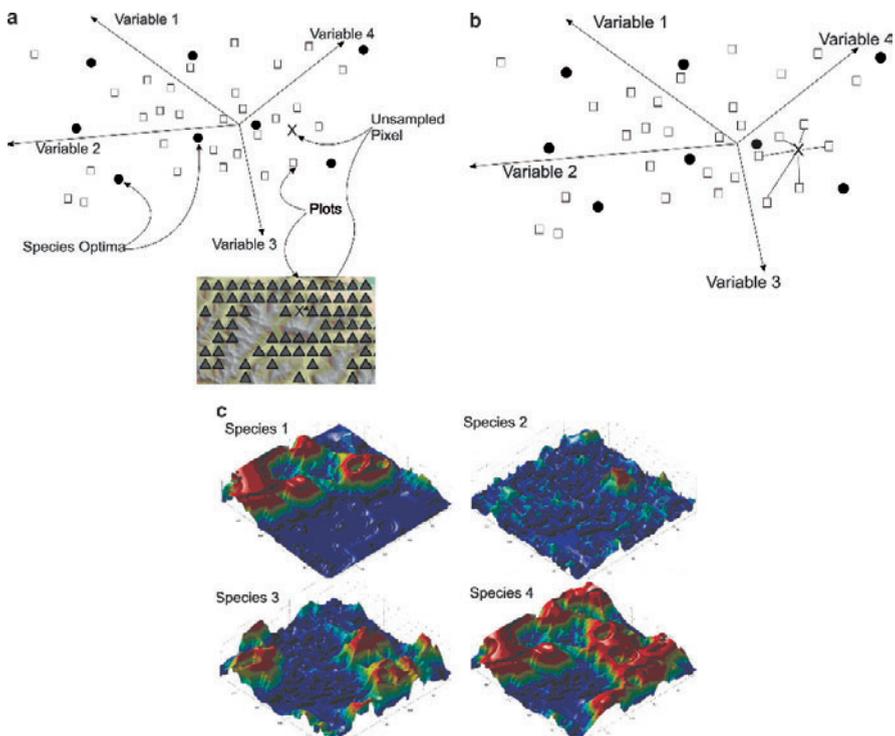
Fig. 6.5 The gradient models produced from the three-dimensional spatial data-cube can be used to produce predictive maps for each response variable through imputation mapping. (**a**) projecting unsampled locations into a gradient model, (**b**) imputing expected values for the unsampled location through imputation, (**c**) products are spatially synoptic predictions of each response variable
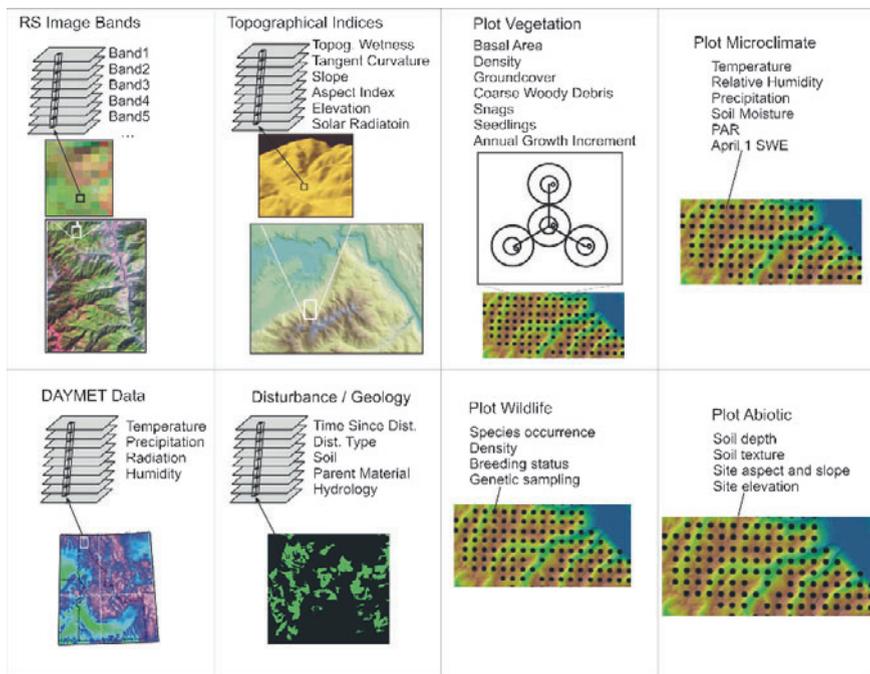
Fig. 6.6 The rich data which comprise the three-dimensional data-cube allow for integration of multiple ecological, management, abiotic and biotic predictiors within a coherent space–time framework for monitoring and prediction

In this approach, plot data, (FIA is an example) in conjunction with unclassified remotely sensed data, DEM, and other spatial data (Fig. 6.6) are used to develop statistical relationships to resources of interest (Fig. 6.4). These algorithms are then used to evaluate resource conditions at the plots (again like FIA) and can be used to produce maps depicting resource condition at fine spatial grain across large analysis areas (Fig. 6.5). In both cases, the reliability of products is directly computable. There are two keys to success (1) extensive and current samples of the condition of multiple resources at many locations across the analysis area, (2) a comprehensive and continually updated geospatial database containing spatial layers describing physiography, disturbance history, and other biophysical attributes, as well as radiometrically corrected bands from multiple remote sensing platforms.

This approach has several advantages over traditional approaches. First, it is fast. Maps are created in near real-time by transforming resource conditions into maps when they are needed. If, for instance, a resource is dependent on the proximity of roads, as soon as the road layer is updated the resource map also automatically updates. Importantly, unclassified satellite data are continually updated, so that changes in forest conditions due to harvest, fire, and other disturbances are immediately reflected in status of all modeled resources. Second, updating is relatively

in expensive. Updating maps is in expensive – for many resources it is automatic and free. Developing the algorithms to produce the maps requires expenditure; applying these algorithms to a shifting landscape does not. Third, it uses fine grain, multivariate base data. The produced maps are therefore always custom designed for the resource of interest, and produced optimally from available multi-scale data. Fourth, it is maximally efficient. Because the approach produces optimized predictions from a large collection of multi-scaled, spatially referenced data layers, mapped descriptions of each modeled resource are the best that can be produced given the base data. Rather than there being a map that is optimal for one purpose and sub-optimal for all other purposes, each purpose has its own independent and optimal map.

A second advantage is localization. Conditions at very fine spatial scales can be inferred through imputation (Ohmann and Gregory 2002; Cushman et al. 2007) or other statistical interpolation approaches. In imputation, relationships between measured plot data, which are scattered spatially, and continuous spatial data, which therefore are available at fine grain sizes for all locations, are related through models to provide estimates of environmental conditions for virtually any desired spatial scale. The power of this approach lies in its ability to use data from the entire extent of measured plots to impute the resource conditions at any location. These predictive models provide rigorous assessment of relationships between the condition of each resource and environmental characteristics at multiple spatial scales and management actions. They further provide predictive maps of resource condition synoptically across space. This provides a means to identify the most likely areas to look for particular rare plant or animal species. It also provides a means to monitor changes in habitat over time, and to predict future changes in habitat amounts and qualities under alternative management and disturbance scenarios.

Gradient modeling can also be used to relate many other point-based data types to continuous data surfaces. For example, gene flow and population connectivity (Cushman et al. 2006; Holderegger and Wagner 2008; Balkhenol et al. 2008) can be modelled using similar methods providing the ability to identify corridors, if they are relevant, barriers, and the factors that influence gene flow for individual species. As described in more detail in Chapter 17, genetic gradient modeling is ideally suited to evaluating the factors that determine connectivity for a range of organisms. For example, it is an ideal approach to determine the factors that drive the spread of different invasive species and identifying the locations most likely to contain incipient populations of invasive species. Identifying these incipient invasions before they become entrenched is critical to effective control of the spread of invasive species. Genetic gradient modeling is also an extremely valuable approach to studying the connectivity of aquatic ecosystems for native and non-native fish and identifying the features that influence connectivity for each species. Furthermore, genetic gradient modeling provides a means of optimizing models of gene flow for a wide range of terrestrial wildlife, providing managers with detailed information to minimize the negative effects of management actions on population isolation and fragmentation if that is determined to be a problem, of species of concern and species of interest.

6.4 Summary and Conclusion

By linking empirical sampling of multiple resources, extensive geospatial databases and sophisticated spatial modeling, integrated resource monitoring has potential to provide estimates of current conditions, measure changes over time, and provide explanation of causes of observed changes, and predictions of expected future conditions. The approach avoids several major assumptions of the coarse filter. First, categorical patch mapping as proxy for habitat is avoided by representing ecosystem diversity as continuously varying gradients of vegetation composition, structure and biophysical variables rather than arbitrary patch mosaics (McGarigal and Cushman 2005; McGarigal et al. in press). Second, it does not assume that a single mapping of patches is an optimal surrogate of habitat for all species. Instead, each resource is predicted individualistically in response to key variables describing multiple biological and abiotic attributes across complex landscapes and over time. Third, it avoids the assumption that the scale of a particular mapped patch mosaic is ideal for all species. Instead, we relate individual species to individual driving variables at a range of spatial scales. Fourth, it provides timely updates of resource maps, using current data to predict resource distribution and condition across space. Fifth, combining gradient modeling with landscape simulation models (Cushman et al. 2007) provides a means to predict distribution and abundance in the future under altered climate, disturbance regime and management.

References

Andelman SJ, Fagan WF (2000) Umbrellas and flagships: efficient conservation surrogates or expensive mistakes? Proc Natl Acad Sci 97:5954–5959

Balkenhol N, Gugerli F, Cushman S, Waits L, Coulon A, Arntzen J, Holderegger R, Wagner H (2007) Identifying future research needs in landscape genetics: where to from here? Landscape Ecology 24:455–463

Block WM, Brennan LA, Gutierrez, RJ (1987) Evaluation of guild-indicator species for use in resource management. Environ Manag 11:265–269

Cushman SA, McGarigal K (2002) Hierarchical, multi-scale decomposition of species-environment relationships. Landsc Ecol 17:637–646

Cushman SA, McGarigal K (2004) Hierarchical analysis of forest bird species-environment relationships in the Oregon Coast Range. Ecological Applications. 14(4):1090–1105.

Cushman SA, Evans J, McGarigal K (in press) Do classified vegetation maps predict the composition of plant communities? The need for Gleasonian landscape ecology. Landsc Ecol

Cushman SA, O'Daugerty E, Ruggiero L (2003) Lanscape-level patterns of avian diversity in the Oregon Coast Range. Ecol Monogr 73:259–281

Cushman SA, O'Daugerty E, Ruggiero L (2004) Hierarchial analysis of forest bird species-environment relationships in the Oregon Coast Range. Ecol Appl 14:1090–1105

Cushman SA, McKelvey KS, Hayden J, Schwartz MK (2006) Gene-flow in complex landscapes: testing multiple hypotheses with causal modeling. Am Nat 168:486–499

Cushman SA, McKenzie D, Peterson DL, Littell J, McKelvey KS (2007) Research agenda for integrated landscape modelling. USDA For Serv Gen Tech Rep RMRS-GTR–194

Cushman SA, O'Daugerty E, Ruggiero L, McKelvey KS, Flather CH, McGarigal K (2008) Do forest community types provide a sufficient basis to evaluate biological diversity? Frontiers in ecology and the environment. 6:13–17

Cushman SA, O'Daugerty E, Ruggiero L (submitted) Sensitivity of habitat and species surrogacies to landscape disturbance. Ecology

Gause GF (1934) The struggle for existence. Williams & Wilkins, Baltimore, MD

Hansen AJ, Urban DL (1992) Avian response to landscape pattern: the role of species' life histories. Landsc Ecol 7:163–180

Holderegger R, Wagner HH (2008) Landscape genetics. Bioscience 58:199–208

Hutchinson GE (1957) Concluding remarks. Cold Spring Harbor Symp Quant Biol 22:414–427

Lambeck RJ (1997) Focal species: a multi-species umbrella for nature conservation. Conservat Biol 11:849–856

Landres PB, Verner J, Thomas JW (1988) Critique of vertebrate indicator species. Conservat Biol 2:316–328

Lindenmayer DB, Manning AD, Smith PL, Possingham HP, Fischer J, Oliver I, McCarthy MA (2002) The focal-species approach and landscape restoration: a critique. Conservat Biol 16:338–345

MacArthur RH (1967) Limiting similarity, convergence and divergence of coexisting species. Am Nat 101:338–345

McGarigal K, Cushman SA (2005) The gradient concept of landscape structure. In: Wiens JA, Moss MR (eds) Issues and perspectives in landscape ecology. Cambridge University Press, Cambridge

McGarigal K, Tagil S, Cushman SA (in press) Surface metrics: an alternative to patch metrics for the quantification of landscape structure. Landsc Ecol

Millspaugh JJ, Thompson III FR, editors (2009) Models for Planning Wildlife Conservation in Large Landscapes. Elsevier Science, San Diego, California, USA. 674 pages.

Noon BR, McKelvey KS, Dickson, BG (2008) Multispecies conservation planning on U S federal lands. In: Millspaugh JJ, Thompson FR (eds) Models for planning wildlife conservation in large landscapes.

Ohmann JL, Gregory MJ (2002) Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in coastal Oregon, USA. Can J Forest Res 32:725–741

Pulliam HR (2000) On the relationship between niche and distribution. Ecol Lett 3:349–361

Roberge J-M, Angelstam P (2004) Usefulness of the umbrella species concept as a conservation tool. Conservat Biol 18:76–85

Schwartz MK, Luikart G, Waples RS, et al (2007) Genetic monitoring as a promising tool for conservation and management. Trends Ecol Evol 22: 25–33

Sherry TW (1979) Competitive interactions and adaptive strategies of American redstarts and least flycatchers in a northern hardwood forest. Auk 96:265–283

Verner J (1984). The guild concept applied to management of bird populations. Environ Manag 8:1–14

Wiens JA., Hayward GD, Holthausen RS, Wisdom MJ (2008) Using surrogate species and groups for conservation planning and management. Bioscience 58:241–252