Chapter 1
Introduction: Ecological Knowledge, Theory and Information in Space and Time

Samuel A. Cushman and Falk Huettmann

A central theme of this book is that there is a strong mutual dependence between explanatory theory, available data and analytical method in determining the lurching progress of ecological knowledge (Fig. 1.1). The two central arguments are first that limits in each of theory, data and method have continuously constrained advances in understanding ecological systems and second that recent revolutionary advances in data and method are enabling unprecedented expansion of ecological investigation into areas of inquiry previously unapproachable due to lack of fine-detail, broad scale data on environmental conditions, the distribution and performance of organisms, the lack of sufficient computational power to process and analyze such voluminous data sets, and inadequate analytical tools to investigate pattern–process relationships among many interacting entities over large, spatially complex landscapes.

1.1 Mutual Dependence of Theory, Method, Data

There is a strict interdependence in science between theory, method and data. It is not possible to decouple these in the practice of science. In some sense it would be desirable if one could. When each corner of this triangle (Fig. 1.1) is dependent and limited by the others there is a feedback where the limitations of each further limit progress in the others. If these could be decoupled conceptually it would perhaps improve the rate of scientific advance. Classic conceptions of the scientific method

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

F. Huettmann
EWHALE lab- Biology and Wildlife Department, Institute of Arctic Biology, University of Alaska-Fairbanks, 419 IRVING I, Fairbanks, AK 99775-7000, USA
e-mail: fffh@uaf.edu

S.A. Cushman and F. Huettmann (eds.), *Spatial Complexity, Informatics, and Wildlife Conservation*
DOI 10.1007/978-4-431-87771-4_1,

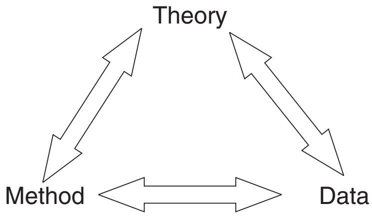


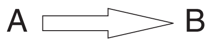
Fig. 1.1 There is a mutual interdependence between methods of observation and analysis, kind and character of data collected, and theories used to explain phenomena. Importantly, there is no possible decoupling by which they will be independent, lending a kind of circularity to the logic of scientific justification. Sometimes, this promotes a self-confirmatory process, with a theory proposing a method designed to produce data that will confirm the theory

Fig. 1.2 Baconian process of induction by which data are collected "objectively", which then suggest appropriate methods for analysis and interpretation, which then suggest the correct theory for explanation

typically assume one of two decouplings. First, the "Baconian" inductive approach proposes a decoupling in which observations (data) are a reflection of reality uncontaminated by implicit theories and unaffected by methods of data collection, such as sampling and measuring (Fig. 1.2). In this conception, the mind, its preconceptions and biases is seen as an obstacle to true understanding and instead the scientist collects data dispassionately and then theory regarding causation emerges inductively from the observations (Bacon 1620). In contrast, the "Cartesian" approach proposes a converse decoupling in which a sentient observer imagines processes governing ideal systems (Descartes 1637). In this conception, it is observation that is unreliable and ideal and eternal conceptions of theory are truly reliable (Fig. 1.3).

Each of these decouplings between theory, method and data are easily refuted. Despite the fact that they are over 400 years old, neither is a realistic view of any actual process used by a practicing scientist to link method, data and theory to build understanding. In the former case, it is easily argued that observations are always "infected" by implicit theory and affected by methods of sampling and measuring. Therefore, it is virtually impossible to obtain purely objective data from which to induce generalizable theory. In addition, due to the logical fallacy of affirming the consequent, patterns observed through induction do not provide proof for a theory with which they may be consistent (Fig. 1.4). The latter case assumes theories are created by the mind, independently from the historical context of current and past explanation. They would be unaffected by the scope and limits of available empirical observations related to the entities and processes related to the theory or by the methods of measurement and analysis that these data are customarily subjected to. These seem severe and unjustifiable assumptions. In addition, strict Cartesian distrust for observation makes empirical evaluation of theory difficult.

Fig. 1.3 Cartesian process of deduction through which a theory is proposed a priori. In strict Cartesian argument, the process stops there, as an ideal theory is seen as superior to the noisy and imperfect methods and data of actual fact. In practice, methods are selected to evaluate the theory, data is collected with these methods, and then used to verify the theory



AFFIRMING THE CONSEQUENT	DENYING THE ANTECEDANT
If A is true then B is true B is true Therefore A is Indeterminate	If A is true then B is true A is not true Therefore B is Indeterminate
MODUS PONENS If A is true then B is true A is true Therefore B is certainly true	MODUS TOLLENS If A is true then B is true B is not true Therefore A is certainly not true

Fig. 1.4 Four logical syllogisms central to scientific reasoning. All four have the same major premise, if A is true then B is true. The bottom two, modus ponens and modus tollens, are logically correct determinate judgments, while the top two, affirming the consequent and denying the antecedent, are indeterminate. Affirming the consequent has special prominence in scientific reasoning and is an abiding challenge to obtaining reliable knowledge

A major focus of this book will be on reasoning within the practice of science, particularly in regard to common logical errors that lead to incorrect conclusions. Figure 1.4 lists four major forms of logical argument from a major and a conditional premise. Two of these (Modus ponens and Modus tolens) are logically correct; the other two (Affirming the Consequent and Denying the Antecedent) lead to logical indeterminacy. A graphical depiction of these four conditional syllogisms may be helpful in gaining an intuitive understanding of the subtle logical traps of affirming the consequent and denying the antecedent (Fig. 1.5). The area outside oval X is –X and the area outside oval Y is –Y. Elements l, m, n are particular members of X, Y, and –Y. In Fig. 1.5, the statement "If X then Y" is represented by the Venn diagram in that all elements of X are also included in Y, so if X is true then Y is also true by overlap, and is a form of the Modus ponens argument. Likewise, the statement "If Y is not true then X is also not true" is represented in that no elements of X exist outside of Y, and is a form of the Modus tolens argument. In contrast, the statement "If Y then X" is not correct in this case, as there are elements of Y (for example m)

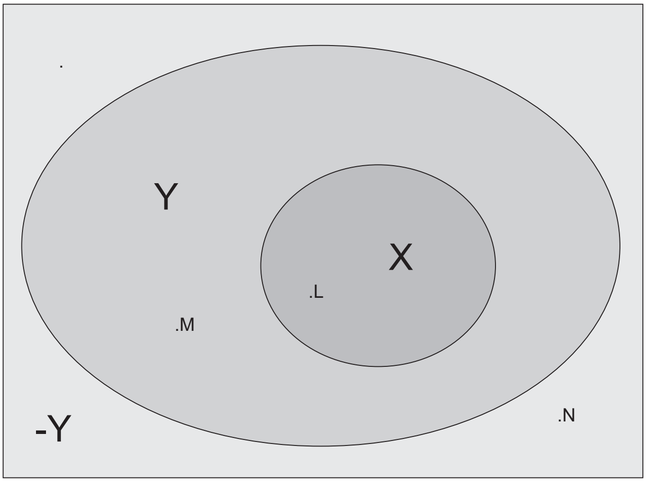


Fig. 1.5 Venn diagram schematic showing the logic behind the four syllogisms listed in Fig. 1.4

that are not also elements of X. This is a case of affirming the consequent, because if X is true then Y would also be true. But it is possible for Y to be true without X being true, as there are elements of Y which do not overlap X. In addition, the conclusion that if X isn't true then Y also isn't true is an example of denying the antecedent, as there are elements of Y which exist outside of X.

One of the inferential challenges of scientific research is that the researcher does not have knowledge of the true relationships between premises a priori. That is, the Venn diagram for a given ecological research question corresponding to Fig. 1.5 is hidden to the observer. That is one reason why data mining and related methods become so popular (Breiman 2001), and model selection so intensely debated (Burnham and Anderson 1998). In deductive reasoning the researcher proposes conceptual Venn diagrams corresponding to hypothetical relationships between causal and response factors. Then observations may be compared to the conceptual model and if they do not match the model may be rejected. In inductive reasoning, the researcher collects data, building evidence, in an effort to confirm universal conditions from a collection of consistent particular observations.

The overall goal of this introductory chapter is to try to link the data–method–theory interdependency (Fig. 1.1) to these four forms of argument within the context of ecological reasoning. The remainder of the book is focused on the details of ecological methods, data and theory, especially in regard to where errors of affirming the consequent and denying the antecedent commonly are made. No definite conclusion can be drawn when the antecedent is denied or the consequent is affirmed. In inferential science both of these errors are extremely common, both when proceeding in an inductive path from data to infer theory, or from a deductive path from theory to confirming data.

The celebrated, so-called "hypothetico-deductive" approach is a modern attempt to partly reconcile deductive and inductive approaches and through a partial fusion

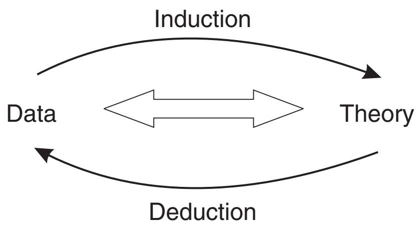


Fig. 1.6 The relationship between data and theory and the pathways of inference of induction and deduction. In induction, theories are induced from a collection of observations, while in deduction a theory is proposed and then data is compared to the theory to see if the theory is consistent or can be falsified

remedy their respective weaknesses (Fig. 1.6). The basic idea is that data without theory are uninformative, while theory without data is not compelling. A scientist proposes a hypothesis to explain some phenomenon, develops a crucial experiment to test it, with the critical attribute that the outcome of the experiment can produce results inconsistent with the hypothesis, and thereby lead to its rejection. The scientist then collects data required by the experiment, conducts the experiment, then either rejects or fails to reject the hypothesis. We believe that the dominant, if often implicit and subconscious, mode of investigation and explanation in modern western science is based on this idea, as elaborated by Karl Popper (1959, 1963) and others, and discussed in a famous paper by Platt published in Science in 1964. While the approach seems to reconcile Baconian induction with Cartesian deduction by linking data to theory through the crucial experiment and refutation, several issues remain. An obvious one is the validity of crucial experiments and falsification.

The criterion of falsification relates to how well defined the research question is and how meaningful is the proposed theory. For example, suppose there are four theories for temperature, T1, T2, T3, T4. The application of these four theories to a particular location and time imply four implications: (I1) Someday the temperature will be above freezing, (I2) tomorrow the temperature will be above freezing, if the conditions are right, (I3) tomorrow the temperature will be above freezing, (I4) tomorrow the temperature will reach at least 20°C. If the temperature tomorrow reaches 18° we will reject theory T4, but not T1–T3. Both T3 and T4 are well-defined statements, in that they stipulate unambiguous criteria for rejecting the theory. However, this is not the case for T1 and T2. Without an independent specification of the necessary and sufficient "if the conditions are right", I2 is true by definition and hence is circular. Also, I1 cannot be falsified because it implies an indefinite period for the observation to occur. Therefore, T1 and T2 are not testable theories, resulting in dubious statements and actions if they should be used as a base for management actions (Hillborn 1997).

An ecologist proposes a theory T for some phenomenon, and identifies and implication I. This then produces a major premise of the form: "if theory T, then implication I." The outcome of the research is data (D), which supplies the minor premise of the conditional syllogism proposed by the theory. Scientists face the indeterminacy of affirming the consequent when D is like I. This is the reason why scientists cannot claim to prove a theory is true even when data consistent with the implication of the theory are obtained. In contrast, when D is inconsistent with I the research conclusion, based on modus tollens, is the rejection of theory T.

The ecologist traditionally has taken a refutationist approach to advancing knowledge. As discussed above there is no means to logically prove a theory through refutation. Thus, empirical science is always in the realm of affirming the consequent. Instead of proving theory, the traditional refutationist paradigm is based on "theory corroboration", through which a researcher is expected to make a concerted and repeated effort to refute the theory, with confidence in the theory supposedly increasing with additional failed efforts to falsify the theory. In this view, science is a process of eternal skepticism, built around the endless quest to disprove theories. This approach in ecology is represented by inferential statistics based on the null hypothesis. The null hypothesis is the implication expressed in data if the theory is not true. It usually is expressed as no difference between treatments in an ANOVA framework or no slope in a regression framework (Hillborn 1997; Zar 1996).

There are at least six major potential criticisms of this "theory corroboration" paradigm using inferential statistics and null hypotheses. First, one may question whether the null hypothesis is a meaningful statement, in terms of the criterion of falsification, described above. We argued that for a theory to be meaningful it must make specific predictions that can be disproved. The traditional null hypothesis has a very weak criterion for falsification, as it is usually expressed in the broadest terms of difference from randomness. Observing statistically significant difference from randomness is a very weak platform to build theory corroboration. This is particularly true for the life sciences, in which true random processes are rare, and there exists a deep foundation of accumulated knowledge of structure–function relationships. Informative hypothesis should at least be based on the latter (O'Connor 2000).

Second, most inferential statistical analysis using null hypotheses has focused on Type I error, or rejecting the null hypothesis when it is true, and not considered Type II errors nearly as carefully. Type II errors are failure to reject the null hypothesis when it is false, and is expressed as rate beta. One minus beta is power, or the probability that one will correctly reject the null hypothesis when it is false. In practice, given the tradeoff between the two error types, the traditional focus on Type I error, and the expense of collecting representative sample sizes large enough to obtain low rates of both error types, actual statistical power is often quite low. With low statistical power the researcher will be unlikely to reject a null hypothesis even if it is false. This clearly cripples one's ability to refute hypotheses, and seriously undercuts the confidence one should have in "corroboration" through repeatedly failing to reject a null hypothesis.

Third, one must question the dedication of practicing scientists to negative inference and skepticism. It's well known that negative results are undervalued and difficult to publish; publication decisions are clearly biased towards positive findings. However, logically, the refutationist approach only advances through rejection of null hypotheses, and does not confirm explanation. However, in practice researchers often treat significant statistical tests as proof of the veracity of the theory. Even when they tacitly admit falsification does not prove a hypothesis, researchers rarely report comprehensive results from many different tests of the same hypothesis to build a measure of theory corroboration. This in large part is due to the nature of scientific publishing in peer-reviewed journals, where a premium is given to space and speed – there is not the space for comprehensive reporting of multiple efforts to refute hypotheses with complementary analyses, nor is there time given the competitive nature of science for researchers to conduct such large, and complex analyses. An extension of this argument would be that even if a researcher were to diligently test a theory against a null hypothesis a number of times with complementary methods and extensive independent data, to what degree does this increase our confidence in the result? One of the insidious attributes of affirming the consequent is that data measured in many ways in many systems may appear consistent with an incorrect theory. There are potentially infinite theories to explain any phenomenon, and a large portion of these might be expected to produce implications measured in data that are equivocal. Thus there is an under-determination of theories by facts (Quine 1953). The data might consistently be consistent with multiple mutually exclusive alternative theories. If this is the case then no amount of "theory corroboration" through failure to refute a null hypothesis really increases our confidence in the theory. We present an example of this in Chapter 17 on landscape genetics.

Fourth, is the challenge of obtaining large, representative samples of ecological data from spatially and temporally complex systems. McGarigal and Cushman (2002) conducted an extensive review of studies to quantify the ecological effects of habitat fragmentation. They noted a dramatic paucity of replicated and controlled studies in which the units of observation were appropriate for landscape-level inferences about pattern–process relationships. The vast majority of published studies used mensurative rather than manipulative study designs focused on relationships between habitat patch size and shape and ecological process. Very few studies combined manipulative design, replication, statistical control, with inference to landscape-level pattern–process relationships. One of the major themes of this book is that spatial and temporal context are fundamentally important to pattern process relationships. The severe shortage of field studies that provide statistically powerful replicated and controlled data sets on ecological responses to spatial and temporal variability at appropriate scales is a major limitation to current knowledge.

Fifth, one of the main themes of this book, which we begin to develop in the following chapter, is that spatial and temporal variability fundamentally alter pattern process relationships such that these are not stationary across space and time. We argue in the following chapters that this implies that research should focus directly on mechanisim–response relationships in context and in particulate. This has major implications for traditional inferential refutationist analysis. For example, if spatial and temporal variability are fundamental factors it will be difficult or impossible to obtain representative and independent samples for inferential analysis. Inferential statistics is based on being able to draw independent and representative observations from a population. This implicitly assumes an urn-model kind of population, where a random draw will produce a representative and independent observation. However if spatial and temporal variability are fundamental we will not be able to obtain either statistical independence or representativeness through random sampling in space. More importantly, if ecological processes are nonstationary in space and time then it will be difficult to define a meaningful population from which to sample. The entire principle of random sampling from a population presupposes that there is a discretely bounded and internally stationary population to sample. If pattern process relationships vary continuously across space and through time, as we will argue throughout this book, then there is no objectively defined population from which to draw independent and representative samples. Instead, as we argue in Chapters 2, 3 and 4, analysis may have to abandon the notion of representative sampling from populations and explicitly adopt a model in which the time-space location of each sample in a large network is explicitly included in the analysis. This poses a tremendous challenge to inferential, population-based, null-hypothesis refutation, a challenge that is not substantially lessened by including time in a repeated measures design or space in a mixed-model form. In addition, linear analysis approaches often cannot cope with these data situations (Craig and Huettmann 2008).

Sixth, naïve falsification has been questioned on several other philosophical grounds. Popper (1963) noted that nearly any statement can be made to fit the data, with the addition of ad hoc adjustments. In addition, the so-called Quine–Duhem thesis holds that for any given set of observations there are an innumerably large number of explanations. Specifically, Quine's (1951) doctrine of confirmation holism emphasizes the under-determination of theories by data, noting that because of the principle of affirming the consequent there may be many equally justifiable alternative explanations for a given observation set. Thus empirical evidence cannot force the revision of a theory. As such, the Quine–Duhem thesis is seen as a refutation of Popper's criterion of falsification as a reliable means of obtaining knowledge.

In recognition of the logical and practical insufficiency of refutationist methods, an alternative paradigm of confirmation has developed. In this approach, instead of proposing a single theory and then seeking disproof, the scientist proposes multiple alternative hypotheses and then seeks to evaluate which explanation is most likely. The exchange is of the questionable logic of critical experiments and decisive refutation for weighing the relative support for multiple alternatives. Likelihood theory, Bayesian belief and the Information Theoretic Approach are all examples of this alternative paradigm. While this appears to circumvent the dilemma of refutation, it embraces an equally challenging dilemma of confirmation. Specifically, there is no logical means to evaluate to what degree the most likely explanation evaluated in a confirmationist approach conforms with any true underlying process. There is just as much inherent risk of affirming the consequent or denying the antecedent as in the refutationist case. Observing that one theory is more consistent with a set of data than another does not prove that theory is correct. As discussed above, there may be any number of alternative theories that could explain the data as well or better (e.g. Quine 1951).

In our view, neither the refutationist nor the confirmationist models of scientific inference address the fundamental interdependence between theory, method and data. We feel that much progress has been made in advancing knowledge using both approaches; yet we also feel both approaches are contingent and particular, rather than providing universal, necessary and certain knowledge of underlying cause and effect. A scientist stands embedded in history, on the shoulders of giants. From that vantage, the view of the scientific landscape, including what questions to ask, on what topics, what methods to use, and likely explanatory hypotheses, are all affected in profound ways by the context of accepted knowledge, prevailing theory, and predominant scientific methods and techniques. Within that context, the prospect for reliably identifying correct relationships between cause and effect is daunting. If spatial and temporal variability are fundamental, and if pattern-process relationships are dependent on them in highly scale-dependent ways, then there may be a strong case for an entirely new approach to ecological inference based on integrating contextual pattern–process relationships explicitly across space and through time, as we discuss in the following several chapters.

There is no possible decoupling of theory, method and data. There is no data that is not infected by theory. There is no theory that is not built on observation. There is no observation that is not dependent on method. There is no method that is not informed and chosen in the context of theory. In the absence of decoupling of method, data and theory, it is difficult to avoid a boring-in of tradition, where theory proposes methods, which produce data, which support the theory, in an endless, inward spiral of self-confirmation. This of course is antithetical to the logical arguments about the risk of affirming the consequent presented above. The most important tool a scientist can possess is skepticism, and the most essential skill is to dispassionately direct this skepticism at one's own theories and observations. This also is the most difficult psychological challenge a scientist faces. Scientists' work is evaluated by their peers within a community which shares common paradigms and assumptions. Thus, even should a scientist have the temerity to challenge the basic assumptions of the field it is very likely their work will receive harsh criticism, while less controversial and more conventional work will usually receive positive reviews as "confirmation" of established belief. This perhaps is the central message of Kuhn's Structure of Scientific Revolutions (1970), that while we stand on the shoulders of giants, we are also shackled by them to triads of self-confirmatory, interdependent traditions of theory, method and data. Thus, there truly is a central role of revolution and paradigm shift in science. Self-confirmatory, boring-in of theory–method–data–theory produces a detailed exegesis of the products of the underlying assumptions of that particular tradition of theory–method–data. But, as in a debate about how many angels can dance on the head of a pin, often further progress in expanding and advancing knowledge of nature depends on breaking out of that inward spiral through questioning assumptions, proposing radical new theories, or developing new methods to obtain novel data and produce new kinds of analyses.

This is the core focus of this book. We believe that the confluence of unprecedented increases in computational power and profound improvements in the ability to measure and sample attributes of the environment at fine scales, across vast spatial expanses and overtime are providing revolutionary changes to two corners of this conceptual triangle. In our conceptual model, this coming revolution in ecological science is being pushed by immense changes in method, specifically computational

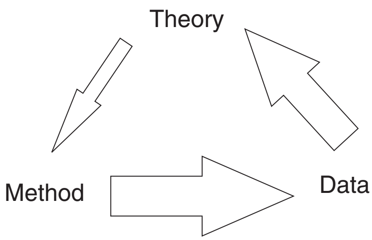


Fig. 1.7 Modern technological advances have enabled revolutionary changes in the kinds and extent of data that can be collected. This in turn has led to enormous advances in methodologies needed to handle, store and analyze these data. This in turn is providing a "bottom-up" forcing on reexamining ecological theory within the context of these new methodologies and data, with new methodlogical advances the dominant driver of change and progress. In contrast, classic ecology of the twentieth century was led by theory, with method and data trailing after. This is because of the combination of the tremendous success of simple and parsimonious mathematical models creating internally consistent explanations, with the lack of ability to collect and analyze large data sets across big spatial extents and over time

power of fast computers, advanced spatial and multivariate statistical and modeling approaches, geographic information systems, remote sensing, and landscape ecology (Fig. 1.7). These, in turn, are vastly altering the amount, kind and scope of data available for analysis, and the kinds of analyses that are possible. At this point in time we are in the birth-throes of new paradigms, with method and data far surpassing theory. The goal of this book is to step through each corner of this triad, to discuss the revolutions of measurement, computation, analysis methodologies, to describe the vast new kinds of data these methodologies produce, and finally, and most difficult, look forward and describe the directions into which we believe these revolutions of method and data will enable the expansion of theory.

The history of ecological science in the twentieth century is marked by the emergence of ecology as a quantitative science. This past revolution in ecological science was led by theory, rather than method or data. It was characterized by the development of ideal mathematical models to describe a broad range of ecosystem and population processes. Importantly, these models provided for the first time formal description of ecological processes that enable prediction and investigation of cause and effect relationships. However, like all ideal models, they typically presented a simplified caricature of processes in which complicating factors, such as spatial variability and temporal disequilibrium, were ignored. Implicit in this is the assumption that equilibrium processes acting non-spatially are sufficient to characterize the major attributes of ecological structure and function, and that interactions of space and time could be shunted into the error term as unexplained and unimportant noise. We believe that nearly all classic quantitative ecological theories of the past century share this characteristic, including Lotka–Volterra predatory prey models, theories of habitat compositional analysis based on proportional use, carrying capacity, predator–prey models, maximum sustainable yield (msy), approaches to analyzing sequential animal movement data based on assumptions of the random use of fixed home ranges, landscape ecology of discrete, categorical patch mosaics, genetics of ideal Hardy–Weinberg populations, population dynamics, dispersal, connectivity, and evolution.

In the final two decades of the twentieth century the sufficiency of these ideal models was broadly called into question. The essence of the concern was whether spatial and temporal variability were in fact noise clouding otherwise ideal non-spatial, equilibrium processes, or whether spatial complexity and temporal variability were central to the basic functioning of ecosystem and population processes in ways that require their explicit incorporation into theory and analysis. This upsurge in theoretical concern about spatial complexity and temporal disequilibrium was enabled by the rise in powerful computing, which allowed scientists to address complex processes in unprecedentedly sophisticated ways. One of the main motivations for adopting ideal, non-spatial, equilibrium models in ecology is that they are mathematically tidy and easy to understand and use to make predictions. On the other hand if the contingencies of history and complexities of spatial pattern across scale are fundamentally important, then these simple, tidy mathematical representations of the world do not suffice. Further, it would be very likely that there will not exist tractable, closed-form alternatives that correspond to a spatially complex and temporally disequilibrial world. We believe this requires a change from seeking closed form equations for processes in ideal systems, to understanding the processes in particulate and in context. By in particulate, we mean understanding the process at the level of the individual entities by representing their actions and interactions explicitly. By in context, we mean representing these particulate processes within the spatial and temporal domains. The advance of powerful computing enabled researchers to simulate processes that could not be represented in closed form equations, and investigate how spatial complexity and temporal dynamics influenced the actions and interactions of ecosystem and population processes. This change in method enabled a change in theory. Prior to the ability to simulate processes in spatially complex and temporally variable systems, few scientists ever asked if idealized models were sufficient, and if they did they likely quickly concluded that addressing spatial and temporal complexity in a hyperdimensional predictor space at the at the level of individual entities was intractable and thus not worth spending time and energy contemplating.

1.2 Informatics and Biology: A Very Powerful But Not Fully Used Melding of Disciplines

Informatics has become a well-established modern scientific discipline; it is taught at most major universities and has emerged as a primary framework in public and private sectors alike for efficiently storing, updating and analyzing vast datasets to address challenging research and economic issues. It provides a general digital and quantitative framework and powerful tools for collecting, compiling, processing and analyzing large and complex datasets and models. However, most of the past informatics applications have not dealt with detailed animal biology or ecology.

Traditionally, informatics has had a strictly technical and mathematical focus, with most past applications in the military and industrial contexts. Sustainability questions were virtually ignored and business applications dominated.

The use of GIS started in North America in the early 1960s: Early efforts were motivated primarily by military, but also by agricultural and landuse issues. GIS is a rather complex topic involving geographic projections, digital databases, spatial analysis and the display of data. Given the technical and informatic challenges posed by large, integrated spatial data storage and analysis, development was rather slow, taking more than 15 years until robust and established software platforms became widely available, allowing to spread GIS applications of relevance for animal ecology. The North American and Australian wildlife community played a major role in applying and implementing GIS into the discipline. However, countries in the Eastern Block, China, India, Africa and Central- and South America were not able to contribute much to GIS, enforcing the digital divide that exist to the very day. GIS applications benefited from the Clinton administration's suspension of Selective Availability for Global Positioning Systems (GPS), which allowed public and commercial use of GPS with high accuracies. This greatly facilitated the collection of highly accurate spatial ecological data in the field, which is a foundation for all efforts to integrate space and time into ecological analysis. Specifically, wildlife surveyors realized that spatially referenced digital databases can offer huge advantages for analysis and quick turn-around of results. Because their data is inherently spatial, GIS provides an optimal platform for storing and analyzing spatial ecological data (Huettmann 2004 for an example). Lastly, many museums and natural-history collections world-wide started to digitize their data holdings into databases, and also realized quickly the value when spatial information is available (Graham et al. 2004). These efforts not only benefit from developments of GIS, GPS, the internet and informatics as a whole, but also from geo-referencing tools such as BioGeoMancer (http://www.biogeomancer.org/), which enable plotting of many data points onto a map and with a quantitative measure of accuracy.

It was quickly realized that large and sophisticated spatial databases require extensive documentation to be reliable and interpretable for multiple users. This in turn prompted the development of global formats and standards for Metadata. Metadata started with the library community, then got added a spatial components (Content Standard for Digital Geospatial Metadata CSDGM by the Federal Geographic Data Committee FGDC http://www.fgdc.gov/) and biological components (Biological Data Profile BDP by the National Biological Information Infrastructure NBII www.nbii.com). They now form the absolute foundation for best available online delivery for a global audience to make best use of the offered huge raw data amounts (Huettmann 2005, 2007b), and for a better ecology and management of natural resources.

Another major development in ecological informatics can be traced to the 1992 Rio Convention, which proposed that digital biodiversity information is a global property and heritage, and thus should be administered as a common good by global agencies such as U.N. and its environmental branches (UNEP; Strong 2001). The Global Biodiversity Information Forum (GBIF, http://www.gbif.org/) was a direct outlet of this movement and it represents the hosting facility for any digital biodiversity information world-wide. Such globally networked ecological databases provide a revolutionary opportunity to collect, share and update information on the global environment as well as a means to compile hyper-variate, multi-scale databases essential for robust ecological analysis and modeling (e.g. Chapter 6). Despite major questions of copyright and information use (e.g. Stiglitz 2006), this represents a new and fresh look to overcome the digital divide, at how informatics can serve Earth and its people for global well-being and sustainability (Huettmann 2005). Secondly, it finally moves informatics back to a tool serving human needs foremost, instead of being used as an intimidating tool 'for the sake of a tool', or to contribute to maintaining unequally distributed wealth. Ecoinformatics will only be effective at contributing to ecology and global sustainability when it is understood and utilized by large portions of the global citizenry within a global democratic process.

It is in this philosophical framework that EcoInformatics and similar disciplines are positioned. Foremost, EcoInformatics deals with the entire ecosystem. As outlined earlier, an ecosystem is a very complex, interconnected entity which provides free ecological services, such as clean air, clean water, healthy climate, pollination and agriculture productivity. Entire human societies, and its wealth, rises and falls with ecological services (Taber and Payne 2003). EcoInformatics provides a framework for society to assure that the most optimal solution is reached for humans managing the Earth. An EcoInformatics and subsequent policy that ignores equal access to information for all citizens, and does not promote environmental justice cannot stand (Stiglitz 2006). It is here where BioInformatics, a simplistic combination of traditional informatics and biology, has not yet achieved the promise of global sustainability. Most university biology departments these days are split into a Biochemical section and a traditional Animal/Wildlife and Plant section. Biology as a discipline suffers from this situation tremendously, and so does BioInformatics, which widely defines itself as using Informatics for the use of Genetic analysis, culminating in applications with an industrial focus such as Genbank and data mining for specific genes occupying high-performance computing centers.

However, Biology is a much wider discipline and foremost, and its disciplines like wildlife management hold the answer to global sustainability, well-being, wealth, and subsequently, human happiness (Taber and Payne 2003). Deep frustration about the narrow one-sidedness in the traditional Biology Departments has already led to split from BioInformatics and its underlying philosophy, and create new disciplines such as EcoInformatics, Biodiversity Informatics, Wildlife Informatics, Ocean Informatics, Polar Informatics, Landscape Informatics, Conservation Informatics etc. This follows the still ongoing and rather problematic paradigm of narrow specialization, which has already set us up for major problems in the sciences, our universities, funding agencies and beyond (Taber and Payne 2003). However, it is hoped that it allows for new and more balanced developments and approaches. Hopefully, these 'new' disciplines can join forces, be fully accessible, compatible and interdisciplinary, and thus, form a new global movement, Sustainability Informatics, where Ecology and technological tools, linked with high-performance computing and online initiatives at it's core, allow to reach global sustainability for the sake of human well-being that fully takes nature into account. The use of decision-making tools would present one of such endavours (Huettmann 2004, 2007c). This new call for 'Sustainability Informatics' is extremely relevant, if our approaches are to make sense and assure long-term survival of animals, biodiversity as a whole, habitats and humans alike. Templates for a sound management exist already: Adaptive Management (Walters 1986). However, the current scheme lacks the inclusion of sound Ecology and Informatics. If these three would get combined (Huettmann 2007d) and also could get manifested in our day-to-day teaching efforts to reach the new generation (Huettmann 2007b), it should allow us to make the best possible decisions, using the available tools and sciences.

This rapid advance of method and data collection enabled a reconsideration of many of the basic assumptions of the preceding dominant paradigm, particularly the sufficiency of equilibrium, non-spatial idealized process models and null-hypothesis inferential statistics. In the chapters that follow we will discuss each corner of this triad. We begin in Chapter 2 where we discuss the history, context and change of a foundational concept in ecology, the species–environment relationship and the structure of biological communities. Chapter 3 addresses the fundamentally important and exceptionally challenging task of handling scale and scaling in ecological relationships. In Chapter 4 we outline emergence of the science of landscape ecology and discusses its potential to provide a coherent framework for the development of a new paradigm of spatial and temporal ecology, and the challenges to be overcome in integrating particulate and contextual processes into existing landscape ecological science. Chapter 5 focuses on the ongoing evolution of landscape ecology, particularly in its conceptual expansion to more effectively address the issues discussed in Chapters 2 and 3. In Chapter 6 we present a framework for data collection to provide flexible, multi-variate, multi-scale, spatially synoptic, temporally dynamic databases to feed modeling and analysis of scale dependent ecological processes in spatially complex and temporally variable environments. Thereafter come a number of chapters on particular kinds of data, with reviews of the attributes, scope and limitations of data on organism occurrence, movement and genetics. Next, we focus explicitly on the next corner of the triad, with four chapters discussing a selection of methodological tools to facilitate spatial and temporal analysis of ecological systems, including remote sensing, geographical information systems, database compilation, software and geostatistics. In the following three chapters we then review current spatial and temporal modeling approaches, focusing particularly on landscape genetics and multi-model resistance modeling. We finally present eight examples from current research on terrestrial and marine ecosystems, and close with chapters on the interaction between public policy and scientific progress and a view into the future and discussion of the outlook for rapid progress to expand theory–method–data to most effectively increase our understanding of the ecological world.

This may seem a daunting and overly ambitious project. We are not trying to invent new paradigms of science, but simply to uncover what has already come into being and give it our interpretation. Such interpretations are necessarily limited by the knowledge, perceptiveness and biases of the interpreters. We hope these thoughts and perspectives may be of use in advancing discussion and promoting debate, leading to improved sciences and decision-making. Despite massive losses of ecological services world-wide, this is a fascinating moment in the history of our science, with tremendous advances in data acquisition, computing, and statistics enabling entirely new approaches to research and opening new avenues to understanding. Simultaneously, we face the unprecedented challenges regarding the state of the world, with climate change, population growth, ballooning consumption all contributing to a global biodiversity crisis. This brave new century appears to be destined to be the century of extinction unless there are dramatic changes to the global socio-economic system. The science of ecology must take a leading role in informing the path of these changes, providing knowledge about ecological systems, human impacts on them, and ways to mitigate our impacts on the biosphere. We hope that this book contributes for these noble goals.

References

Bacon F (1620) The New Organon (Cambridge Texts in the History of Philosophy)/Francis Bacon (ISBN 0521564832)

Breiman L (2001) Statistical modelling: the two cultures. Statistical Sci 16: 199–231

Burnham KP, Anderson DR (1998) Model Selection and Inference: a Practical Information-Theoretic Approach. Springer, New York

Craig E, Huettmann F (2008) Using “blackbox” algorithms such as TreeNet and Random Forests for data-mining and for finding meaningful patterns, relationships and outliers in complex ecological data: an overview, an example using golden eagle satellite data and an outlook for a promising future. Chapter IV in: W. Hsiao-fan (Ed.). Intelligent Data Analysis: Developing New Methodologies Through Pattern Discovery and Recovery. IGI Global, Hershey, PA, USA

Czech B. (2008) Prospects for reconciling the conflict between economic growth and biodiversity conservation with technological progress. Cons Bio 22: 1389–1398

Descartes R (1637) A Discourse on method. Pomona Press

Diamond, J (1999) Guns, Germs and Steel: The Fates of Human Societies. Norton, New York

Graham CH, Ferrier S, Huettmann F, Moritz C, Peterson AT (2004) New developments in museum-based informatics and applications in biodiversity analysis. Trends Ecol Evol 19: 497–503

Hillborn R (1997) The Ecological Detective: Confronting Models with Data. Princeton University Press, Princeton, NJ

Huettmann F (2004) Towards a Marine Environmental Information System (MEnvIS) for the Northwest Atlantic: Experiences and suggestions from a multi-disciplinary GIS conservation research project using public large scale monitoring and research data bases from the internet. 4th WSEAS (World Scientific and Engineering Academy and Society) Proceedings, Tenerife/Spain, December 2003

Huettmann F (2005) Databases and science-based management in the context of wildlife and habitat: Towards a certified ISO standard for objective decision-making for the global community by using the internet. J Wildlife Manage 69: 466–472

Huettmann F (2007a) German Environmental Informatics vs. Nature Conservation, GBIF, Rio, Statistics and Sustainability: Are we riding a dead horse? (in German). Rundbrief des Fachausschusses der Umwelt Informatik, Deutsche Gesellschaft fuer Informatik (GI) Nr 42: 9–14

Huettmann F (2007b) The digital teaching legacy of the International Polar Year (IPY): Details of a present to the global village for achieving sustainability, in: M. Tjoa and R.R. Wagner (Eds). Proceedings of the 18th International Workshop on Database and Expert Systems Applications (DEXA). IEEE Computer Society, Los Alamitos, CA, pp. 673–677

Huettmann F (2007c) Constraints, suggested solutions and an outlook towards a new digital culture for the oceans and beyond: experiences from five predictive GIS models that contribute to global management, conservation and study of marine wildlife and habitat, in: E. Vanden Berghe et al. (Eds.). Proceedings of 'Ocean Biodiversity Informatics': An International Conference on Marine Biodiversity Data Management Hamburg, Germany, 29 November–1 December, 2004. IOC Workshop Report, 202, VLIZ Special Publication 37: pp. http://www.vliz.be/vmdcdata/imis2/imis.php?module=ref&refid=107201

Huettmann F (2007d) Modern adaptive management: Adding digital opportunities towards a sustainable world with new values. Forum Public Policy: Climate Change Sustain Dev 3: 337–342

Kuhn TS (1970) The Structure of Scientific Revolutions. University of Chicago Press, Chicago, IL

McGarigal K, Cushman SA (2002) Comparative evaluation of experimental approaches to the study of habitat fragmentation effects. Ecol Appl 12: 335–345

O'Connor R (2000) Why ecology lags behind biology. Scientist 14: 35

Platt JK (1964) Strong inference: Certain systematic methods of scientific thinking may produce much more rapid progress than others. Science 146: 347–353

Popper K (1959) The Logic of Scientific Discovery. Basic Books, New York

Popper K (1963) Conjectures and Refutations. Routledge, London

Quine WVO (1951) Two dogmas of empiricism. Philos Rev 60: 20–43

Quine WVO (1953) From a Logical point of view. Harvard University Press. ISBN 0-674-32351-3

Stiglitz JE (2006) Making Globalization Work. Norton, New York

Strong M (2001) Where on Earth Are We Going ? Texere, New York

Taber RD, Payne NF (2003) Wildlife, Conservation, and Human Welfare: A United States and Canadian Perspective. Krieger, Malabar, FL

Verner J, Morrison ML, Ralph CJ (1984). Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates: Based on an International Symposium Held at Stanford Sierra Camp, Fallen Leaf Lake, CA, 7–11

Walters C (1986) Adaptive Management of Renewable Resources. Blackburn, Caldwell, NJ

Weizenbaum J (1984) Computer Power and Human Reason. Pelican, London

Zar JH (1996) Biostatistical Analysis. 4 ed., Prentice-Hall, Upper Sadle River, NJ