Introduction

Predictive ecology: systems approaches

Matthew R. Evans¹,*, Ken J. Norris² and Tim G. Benton³

¹Centre for Ecology and Conservation, School of Biosciences, University of Exeter, Cornwall Campus, Penryn, Cornwall TR10 9EZ, UK
²Centre for Agri-Environmental Research, School of Agriculture, Policy and Development, University of Reading, Earley Gate, Reading RG6 6AR, UK
³Faculty of Biological Sciences, University of Leeds, Leeds LS2 9JT, UK

The world is experiencing significant, largely anthropogenically induced, environmental change. This will impact on the biological world and we need to be able to forecast its effects. In order to produce such forecasts, ecology needs to become more predictive—to develop the ability to understand how ecological systems will behave in future, changed, conditions. Further development of process-based models is required to allow such predictions to be made. Critical to the development of such models will be achieving a balance between the brute-force approach that naively attempts to include everything, and over simplification that throws out important heterogeneities at various levels. Central to this will be the recognition that individuals are the elementary particles of all ecological systems. As such it will be necessary to understand the effect of evolution on ecological systems, particularly when exposed to environmental change. However, insights from evolutionary biology will help the development of models even when data may be sparse. Process-based models are more common, and are used for forecasting, in other disciplines, e.g. climatology and molecular systems biology. Tools and techniques developed in these endeavours can be appropriated into ecological modelling, but it will also be necessary to develop the science of ecoinformatics along with approaches specific to ecological problems. The impetus for this effort should come from the demand coming from society to understand the effects of environmental change on the world and what might be performed to mitigate or adapt to them.

Keywords: ecological modelling; prediction; climate change; evolution; systems biology; global circulation model

'We appeal to the notorious fact that ZOOLOGY, soon after the commencement of the latter half of the last century, was falling abroad, weighed down and crushed, as it were, by the inordinate number and manifoldness of facts and phenomena apparently separate, without evincing the least promise of systematizing itself by any inward combination, any vital interdependence of its parts'.

Samuel Taylor Coleridge [1]

The ability to predict, to forecast how a system might behave in the future, is a key feature of any science. Prediction is, according to, the Nobel laureate, immunologist and philosopher of science, Peter Medawar, a 'property that sets the genuine sciences apart from those that arrogate to themselves the title without really earning it' [2]. Some subjects, e.g. economics, engineering, climatology, routinely make predictions that others can use; and people are prepared to rely on them as a guide to the future, despite them sometimes, in the light of experience, being inaccurate. It was a premise of the Royal Society discussion meeting, upon which this volume is based, that ecology needs to become more predictive, especially in the context of creating an understanding of the way in which biological systems might respond to environmental change. It is well understood that making predictions about how any system will behave in novel circumstances is difficult. Making predictions about how noisy ecological systems will behave when exposed to large-scale environmental changes will be extremely difficult [3]. Nevertheless, there is a demand for the production of such predictions in order to understand the possible impacts of phenomena such as habitat fragmentation, climate change, invasive species, over-exploitation and pollution on the biological world [4].

Quantification in ecology has historically been focused on the problem of trying to discern the effects of parameters of interest within noisy, real-world systems. Therefore, ecologists have typically been concerned with producing predictions within the observational bounds of their datasets; describing the behaviour of complicated systems has been enough of a challenge. It is our contention that ecologists...
need to embrace the further challenge of developing truly predictive models of their systems, models that can be cast into future, often novel, conditions in which the systems have not yet been observed. We believe that this is necessary to allow ecologists to improve their understanding of the systems they study but is particularly important to enable us to address questions about the impact of environmental change on the biological world. Anthropogenic environmental change is creating novel environmental conditions and we need to develop the ability to forecast the way in which ecosystems, components of ecosystems and the services we derive from them, will alter as conditions change [5]. Although, at least conceptually, this problem could be approached empirically—perhaps by experimentally manipulating a system, this would be extremely difficult at any significant scale; the most likely way to predict the future state of ecological systems will be through modelling.

The first issue encountered in the development of a model is that all models are abstractions; it is obviously impossible to include every aspect of the real world in any model. Once a modeller recognizes that they cannot include all variables in a model, they have to make decisions about which to retain in the model and which to omit [6]. Such decisions will modify the outcomes achieved by a model and the process should be guided by the modellers’ objectives. This, at heart, a philosophical question and although they would not often recognize they are doing so, ecological scientists are constantly making decisions based on a philosophical stance [7]. Model building, in particular, needs to be conducted with an eye to philosophy; for example, whether one believes that a model can be both maximally general (apply to many systems) and realistic (produces predictions that are accurate reflections of a specific system) [7–9], or believes that these two desirable model attributes cannot be simultaneously maximized [10–14], is a philosophical stance. One’s answer to this question should influence one’s opinion about how to proceed in modelling the ecological impact of environmental change. If a realistic model is needed, so that its output can be used to understand a real-life system (i.e. if we wish to use the model to inform us about the fate of a particular system) then, if you believe that generality and realism cannot be simultaneously maximized, you should reject the simple-but-general approach that gives preference to attaining analytical soluble formulations [15], because such models are difficult to use for the purpose of generating realistic answers for specific systems of interest [16].

An important feature of anthropogenic environmental change is that it produces novel conditions. This presents challenges for modelling as statistically derived functions cannot legitimately be projected beyond the bounds of observation, leading one to conclude that phenomenological models should not be projected into novel conditions [16]. If this is the case then neither of the two modelling paradigms that dominate ecology today—simple regression-based models and analytical simplifications containing a small number of parameters—will be particularly useful in the endeavour of making ecological predictions, and so an alternative approach is needed. A systems approach in which characteristics of one level in a hierarchy are explored as emergent properties of processes lower down in the hierarchy [17], will be essential for making ecological predictions in novel conditions. This is because systems approaches do not assume that a description of a system will remain valid indefinitely (as do phenomenological models by definition), they rely on the fact that the internal processes will continue to operate into the future and that their operation will be in some way altered by the changed conditions. The higher order emergent properties alter as a consequence of the changes in the internal processes not because the higher order effects themselves have been projected into the future.

The adoption of a systems approach to the analysis of ecological systems is not novel. Systems ecology was the term used to describe the approach promulgated by the International Biological Programme (IBP) in the 1960s [18,19]. The aim of this approach was to measure as many aspects of the ecological system as feasible and then search for relationships and link data together using the abilities created by the computing power that was then becoming increasingly available [19]. The systems ecology of the 1960s did not develop into a major activity in subsequent years and the funding for the IBP ended in 1974, partly because of the criticism it received on philosophical grounds from other prominent ecologists, particularly Levins [10,12]. However, as succinctly stated by Peter Hudson in the discussion meeting ‘the world has moved on since the 1960s’. At the very least, there has been substantial progress in the scale of computing power; according to Moore’s Law since 1960 computing power has increased by a factor of approximately $5 \times 10^7$, but there has been simultaneous increase in our ability to collect data through both remote and miniaturized technologies. Perhaps even more importantly, behavioural ecology [20] and life-history theory [21] have emerged successfully to provide functional explanations for individual behaviours and decision-making. Therefore, unlike the 1960s, we have explanations for why organisms behave in the way that they do, as well as what they do.

One difference between ecology in 2011 and ecology in the 1960s is that we have increasingly detailed studies on model systems. One of these is the soil mite, Sancassania berlesei, which has been the subject of study for 15 years by Benton and co-workers. A process-based model which incorporated eight ‘genes’ (more correctly allocation rules) that controlled the way in which individual mites use resources displayed how rapidly the signature of evolutionary change can be seen in the population dynamics (number of adults and juveniles both decline over times), life-history theory has increased in our ability to collect data through both remote and miniaturized technologies. Perhaps even more importantly, behavioural ecology [20] and life-history theory [21] have emerged successfully to provide functional explanations for individual behaviours and decision-making. Therefore, unlike the 1960s, we have explanations for why organisms behave in the way that they do, as well as what they do.

A systems approach in which characteristics of one level in a hierarchy are explored as emergent properties of processes lower down in the hierarchy [17], will be essential for making ecological predictions in novel conditions. This is because systems approaches do not assume that a description of a system will remain valid indefinitely (as do phenomenological models by definition), they rely on the fact that the internal processes will continue to operate into the future and that their operation will be in some way altered by the changed conditions. The higher order emergent properties alter as a consequence of the changes in the internal processes not because the higher order effects themselves have been projected into the future.

The adoption of a systems approach to the analysis of ecological systems is not novel. Systems ecology was the term used to describe the approach promulgated by the International Biological Programme (IBP) in the 1960s [18,19]. The aim of this approach was to measure as many aspects of the ecological system as feasible and then search for relationships and link data together using the abilities created by the computing power that was then becoming increasingly available [19]. The systems ecology of the 1960s did not develop into a major activity in subsequent years and the funding for the IBP ended in 1974, partly because of the criticism it received on philosophical grounds from other prominent ecologists, particularly Levins [10,12]. However, as succinctly stated by Peter Hudson in the discussion meeting ‘the world has moved on since the 1960s’. At the very least, there has been substantial progress in the scale of computing power; according to Moore’s Law since 1960 computing power has increased by a factor of approximately $5 \times 10^7$, but there has been simultaneous increase in our ability to collect data through both remote and miniaturized technologies. Perhaps even more importantly, behavioural ecology [20] and life-history theory [21] have emerged successfully to provide functional explanations for individual behaviours and decision-making. Therefore, unlike the 1960s, we have explanations for why organisms behave in the way that they do, as well as what they do.

One difference between ecology in 2011 and ecology in the 1960s is that we have increasingly detailed studies on model systems. One of these is the soil mite, Sancassania berlesei, which has been the subject of study for 15 years by Benton and co-workers. A process-based model which incorporated eight ‘genes’ (more correctly allocation rules) that controlled the way in which individual mites use resources displayed how rapidly the signature of evolutionary change can be seen in the population dynamics (number of adults and juveniles both decline over times), life-history theory has increased in our ability to collect data through both remote and miniaturized technologies. Perhaps even more importantly, behavioural ecology [20] and life-history theory [21] have emerged successfully to provide functional explanations for individual behaviours and decision-making. Therefore, unlike the 1960s, we have explanations for why organisms behave in the way that they do, as well as what they do.
All ecologists will, of course, be aware of evolution but there is a tendency to ignore its effects once one moves beyond the individual level of organization. The prospect of both predicting into the future and that the future into which we might wish to make predictions will be different from the present makes the case for the inclusion of evolutionary change into ecological models.

Including evolution within ecological models is logical if we know that the signature of evolution can be seen in the timescale of our models. But the need to include evolution in ecological models is not simply a matter of ‘if we know it exists then we should include it’. Ecological and evolutionary change are intertwined—population dynamics are the product of the realized life histories (a product of selection) of individuals within the population, whereas the strength of selection is modified by properties of the population (e.g. density) [24]. Including evolution will add a level of complexity, but work in environmental science provides us with tools to start addressing these issues. For example, if we know that organisms act so as to maximize relative fitness [25,26], then we have a conceptual basis for modelling life-history decisions of organisms, even when we have sparse information about their biology. Including evolution may make modelling with limited data more straightforward. This is particularly important given the recognition that individuals experience different environments, make different life-history decisions, and therefore vary in their demographic traits, with this inter-individual variance being important for population and evolutionary dynamics. Further embedding evolution in ecological thinking would allow greater appreciation of the relationship between biological processes at the individual level and the population, community or ecosystem results of these processes. If novel environmental change imposes selection pressures that cause the components of an ecological system to evolve and if this evolutionary process affects the way that system functions, then evolution cannot be ignored.

Despite the desirability of including evolution in ecological models, the most successful group of process-based models in ecology—the terrestrial biosphere models do not do so [27–31]. This may be due to the fact that their raison d’être is to predict changes in the forest community and so even though they have long simulated run times—typically 1000 years, the organisms within them also have long generation times. Long generations relative to model run time, coupled with relatively weak selection pressures that are largely independent of life-history traits (i.e. areas are usually either clear-felled or left alone) may mean that selection can be safely neglected in these models. However, the same would not be true of the populations of herbivorous insects that might live on such trees. A good example of the value of incorporating evolution into ecological models is seen in the analysis of fisheries data. Fishing exerts strong directional selection on fish stocks as it typically removes large adult fish from the population. Fish populations of several species have responded to such selection by maturing earlier and smaller [32,33]. The commonest form of fishing management—imposing a quota that can be filled after which no fish can be landed—appears to result in more rapid evolutionary change than either constant rate harvesting or the maintenance of a constant stock population [34]. These evolutionary changes then translate into more rapid decreases in population biomass in a stock managed by a quota system than in one managed by alternative approaches [34]. The conclusion from this must be that the consequences of evolutionary change on stock abundance and sustainability need to be taken into account when considering management options and that this can be achieved only by incorporating evolution into the ecological models that underpin decision-making in fisheries. In addition, these examples make it clear that there are ecosystem-level consequences of the evolutionary process created by fishing pressure; this feedback clearly needs to be understood before we can understand fully the consequences of the fishing pressure on ecosystem function.

Terrestrial biosphere models show the levels of complexity that can currently be achieved using process-based models. Medvigy and Moorcroft demonstrate that the terrestrial biosphere model Ecosystem Demography 2 (ED2) parametrized using 2 years of data from a single (relatively small) site can make accurate predictions of the patterns of biomass dynamics and community composition over a region of northeastern North America spanning about 15° of longitude and 10° of latitude [35]. This region spans ecosystems that range from almost entirely deciduous forests in the south to conifer-dominated forests in the north and varies hugely in climate forcing and land-use history. The heterogeneity of the output across the region is driven by the differential responses of the five plant functional types within the model. An interesting observation is that ED2’s predictions for certain parameters begins to drop off in the highest latitudes, probably because of the fact that the dominant late-successional tree in this region is not found in the site used for parametrization, raising the possibility that additional functional types may be needed to be added to the model the further it is extended away from the area from which its parameters are derived. ED2 uses size- and age-structured partial differential equations to approximate the first moment behaviour of a stochastic individual-based model [27,36]. The process of using such an approximation makes the model more computationally tractable but means that information is lost about individual heterogeneities. Clark et al. [37] argue that as it is individuals rather than populations that respond to climate through their responses to weather conditions, then it is important to track responses to climate on an individual basis. Vulnerability of a species to climate change can be inferred by aggregating individual-level data [37]. This is a specific example of the general principle that individual organisms are the elementary particles of ecology [38].

The basic level of ecological heterogeneity is found at the level of the individual, although one must recognize that individuals are themselves emergent properties of interactions between genes in a particular environmental context. Inter-individual variation provides the raw material for natural selection (and through the differential success of individuals we get gene-level selection), and is the basis for ecological heterogeneity in both space and time. Models must make simplifying
assumptions in order to make their problems tractable [6,10]. Any process that discards heterogeneity will result in the loss of information about the system, whether this is the selection of functional types rather than species [35], population averaging of survival and growth rates rather than accounting for the importance of individual interactions [37] or assuming that species range may be adequately described by a few climatological parameters [39]. The heterogeneity at different levels of organization and how this affects emergent properties at higher levels of organization needs to be considered. Any loss of heterogeneity is likely to alter the outcome of the model and the key issue is whether it does so in an important manner.

It would be difficult to locate an ecosystem that had escaped modification by humans in some manner. The meeting presupposed the need to consider the ecological impact of anthropogenic environmental changes as a major rationale for the need to make ecological predictions. Clearly, human activity often plays an important role in structuring ecosystems and is frequently a driver of change within ecosystems [40,41]. The exclusion of humans from ecological models seems artificial, and although there are some good examples in which individual- or agent-based models of human populations (agent-based models (ABMs) in the social sciences being essentially similar to individual-based models in ecology) and biosphere models have been successfully coupled to produce predictions about the fate of the ecosystem and consequences to the human population [42], these are unusual. Agent-based modelling is regarded as an appropriate tool for understanding human decision-making and is often used in models of human–environment interactions [40,41]. Rounsevell [40] identifies that many of the issues that concern the ecological community about individual-based models (IBMs) are also of concern to social scientists using ABMs of human systems. For example, processes need to be scaled both out and up to include larger geographical areas (out-scaling) and to aggregate individuals into groups (up-scaling). These would be familiar problems to people concerned with using IBMs and there may be parallel solutions in, for example, the use of stage- and age-structured approximations to scale out [35] and functional types to scale up [40], although any such approximations would inevitably result in a reduction of individual-level heterogeneity in the model, which if it influenced the outcome in important ways might be of concern [37].

Human behaviour is frequently the target of policy interventions both by agencies attempting to create their desired conservation outcomes [17,41] and by institutions and governments [40]. The fact that humans often modify their behaviour extremely rapidly in response to an intervention can mean that there are unintended outcomes of well-intentioned interventions; for example, the fact that acquisition of land for nature conservation can impact upon local land prices potentially results in landowners being tempted by higher prices to put additional land on the market leading to a counter-productive effect on biodiversity outside protected areas if this land is then developed [43]. Milner-Gulland [41] points out that such dynamic behaviour by people results in statistical models only being predictive while the conditions under which the model was created prevail. This is a special case of the inevitable problems associated with using phenomenological models: statistical functions cannot be legitimately projected beyond the bounds of observation [44]. Process-based systems models are therefore needed to understand the dynamics of socio-ecological systems; such models could allow the appropriate feedbacks between levels (e.g. management actions, resource user decisions and ecological sustainability) to be incorporated [41] in much the same way as they are needed to understand the impacts of environmental change on ecosystems, again with feedbacks between levels (e.g. individuals, populations, communities, ecosystem). If we believe that humans are both key to the functioning of many ecosystems (e.g. humans maintain habitats such as grouse moors in their present state) as well as drivers of change in ecosystems (e.g. humans take decisions to change from managing a moor for grouse to forestry), then such models will only be complete when they incorporate both the human and ecological systems.

Process-based models that incorporate realistic amounts of complexity will be computationally and technically demanding. Systems approaches are much more familiar in other disciplines than they are in ecology and it will be possible to acquire skills, techniques and approaches from fields such as climatology and molecular systems biology. Bioinformatics has transformed molecular biology; a combination of the availability of high-throughput data and computing power led to the emergence of molecular systems biology in the last years of the twentieth century [45]. There are now a range of techniques available to molecular biologists who wish to adopt a systems approach to modelling, some of which can be applied to ecological datasets [46]. Bioinformatic techniques were applied to almost 50 years of fisheries data from the Georges Bank, East Scotia Shelf and North Sea to determine the species relevant to the functional collapse of the cod stocks in Georges Bank. The same model was then applied to the other two areas. An interesting result from this analysis suggests that very different species seem to play equivalent roles in the functional collapse of the three ecosystems, namely zooplankton species in Georges Bank, deepwater fish species on the East Scotian Shelf and commercially exploited demersal species in the North Sea [46]. It also suggests that there is a parameter(s) external to the fishery that influences its collapse; it would be difficult to detect such a parameter in conventional analyses of fisheries [46].

In molecular biology, data can emerge from high-throughput machines and the challenge is usually in handling and processing data volumes, but ecological data are usually less abundant and are often less easily sourced. To combat this, a significant advance could be made in ecology if scientists working in the discipline adopted the habit of data-sharing (see http://www.datadryad.org), a habit that prevails in many other areas of science—including molecular biology. As a result of restricted, often private, knowledge of the existence of particular data, it is not always
clear what is known and it is highly likely that the data are available to allow us to know more than we think we know. If data sharing became the norm it is likely that, as in molecular biology, the benefits would far outweigh the costs [47]. The development of the bioinformatics of ecology, ‘eoinformatics’ [46], will be facilitated by the development of new analytical tools as well as by the adoption of tools from elsewhere. Pattern-orientated modelling (POM) may be such an eoinformatic tool. POM concerns itself with identifying a set of patterns that can be observed at different scales and levels and are associated with a real-life problem of interest. Models from which the same patterns emerge are likely to contain the right mechanisms to describe the problem [48].

We should also not forget that some areas of ecology have taken more steps down the path of adopting systems thinking than others. The biosphere models are an example of a process-based approach applied to one trophic level of an ecosystem [27–31,35]. There have been moves in the direction of adding further trophic levels; for example, some similar models include bark beetle-induced mortality [49] and the effect of pine beetle outbreaks on productivity and carbon capture and storage [50]. One sphere of activity in which the use of systems approaches may be inevitable is in the analysis of ecosystem service provision; for example, an analysis of changes in one ecosystem service—soil carbon—required systems thinking because the output was dependent on diverse components of the system, ranging from the climate to anthropogenic land use to biogeochemistry [51]. The consideration of the provision of ecosystem services moves us to the position where we need to consider the values and benefits of those services, rather than traditional arguments about the preservation of iconic species or habitats [17]. This means that we will need to consider more fully the way in which these parameters depend on the changes in biodiversity created by environmental change. This almost inevitably becomes a problem that requires a systems approach, as changes in the value of goods and services emerge from ecosystem processes at various levels [17].

It would be a travesty of the process-based approach advocated here to suggest that it meant that every individual organism in an ecosystem needed to be included in a model of that ecosystem in order to make useful predictions about it. Not only would such a naive exercise be extremely difficult to construct and parametrize, but also it would be almost impossible to comprehend and would probably exceed the ability of even modern computers to run the computations [10]. One challenge for a modern formulation of systems ecology will be to create models of the ‘appropriate’ complexity. The appropriate heuristic is ‘Einstein’s razor’: ‘everything should be as simple as possible, but no simpler’ [52]. Global circulation models (GCMs) are good examples of predictive modelling and have developed both great explanatory and persuasive power. GCMs are process-based in the sense that climate is an emergent property of the processes internal to the system, yet many of those internal processes remain as black boxes and many are recognized as being in need of refinement, e.g. the biosphere–atmosphere feedbacks [53]. Predictive models of ecosystems will need to tread a line between increasing complexity and decreasing comprehensibility; the position along such a line will depend on the needs of the end user. It is possible that if the end user is a policy-maker (in a general sense), then it may be unnecessary to have models with continuous distributions of states—categories may be sufficient and it may be possible to consider the probability of transitions occurring between categories instead of displaying, possibly, artificial levels of precision [54].

If we are to move ecology towards prediction, especially if we are to embrace the need to make ecological forecasts about the impact of environmental changes on the biological world [5], then we will need to further engage with process-based ecological models. We believe that there is a good philosophical basis for doing so, especially if we are concerned with making predictions outside the range of conditions within which data were collected [16], but also if we are concerned with making predictions that are realistic rather than general [16] (see also [7]). We believe that there will be challenges and opportunities in this endeavour; for example, the inclusion of evolutionary change into ecological models will be necessary in a way that is required in neither physical science models nor molecular systems biology [24,55,56]. Evolution can happen surprisingly quickly [22,23] and it will not be reasonable to keep the distinction between the evolutionary play and the ecological theatre [55]. Simplifying assumptions will be necessary to make ecological systems models analytically, and computationally and intellectually tractable. The heterogeneity between individuals matters in ecology, unlike in physicochemical systems (although such systems may be as stochastic as ecological ones). There is heterogeneity at all levels and it can affect the outcome of our processes [35,37]. Ecological models will be made additionally complicated by the need to include humans if, as they often will be, they are important ecological players as well as drivers of ecological change [40,41]. However, socio-ecological modellers already use ABMs that are well adapted to integration with process-based ecological models. In order to help with the endeavour of creating a modern systems ecology, there are techniques and approaches that already exist in other fields and which can be applied to ecological systems [46]. If we wish to consider the creation of an ecological parallel to bioinformatics, then we will need to create new tools and approaches to allow us to cope with the fact that ecology has different constraints and advantages to molecular systems biology [48]. Further developing systems ecology to a point where it can make useful predictions about the ecological impact of environmental change, equivalent to those regularly made by climatologists using GCMs about the future state of the climate, will be demanding. Nevertheless, the need from society to maintain and improve its ecosystem services, all of which derive from ecological systems, should provide sufficient imperative to overcome these complexities [17]. Reaching this point would bring significant benefits; not only will it

Phil. Trans. R. Soc. B (2012)
increase our comprehension of the natural world but it will also help us move towards truly predictive ecology.

REFERENCES


