



The pitfalls of ecological forecasting

TOM H. OLIVER* and DAVID B. ROY

NERC Centre for Ecology and Hydrology, Wallingford, Oxfordshire, OX10 8BB, UK

Received 23 December 2014; revised 21 April 2015; accepted for publication 21 April 2015

Ecological forecasting is difficult but essential, because reactive management results in corrective actions that are often too late to avert significant environmental damage. Here, we appraise different forecasting methods with a particular focus on the modelling of species populations. We show how simple extrapolation of current trends in state is often inadequate because environmental drivers change in intensity over time and new drivers emerge. However, statistical models, incorporating relationships with drivers, simply offset the prediction problem, requiring us to forecast how the drivers will themselves change over time. Some authors approach this problem by focusing in detail on a single driver, whilst others use ‘storyline’ scenarios, which consider projected changes in a wide range of different drivers. We explain why both approaches are problematic and identify a compromise to model key drivers and interactions along with possible response options to help inform environmental management. We also highlight the crucial role of validation of forecasts using independent data. Although these issues are relevant for all types of ecological forecasting, we provide examples based on forecasts for populations of UK butterflies. We show how a high goodness-of-fit for models used to calibrate data is not sufficient for good forecasting. Long-term biological recording schemes rather than experiments will often provide data for ecological forecasting and validation because these schemes allow capture of landscape-scale land-use effects and their interactions with other drivers. © 2015 The Linnean Society of London, *Biological Journal of the Linnean Society*, 2015, 115, 767–778.

ADDITIONAL KEYWORDS: butterfly monitoring – density dependence – environmental drivers – prediction – predictive modelling – weather.

WHY ATTEMPT TO FORECAST?

Two very different masters teach [man] his lesson: experience and foresight. Experience teaches efficiently but brutally...I should prefer, in so far as possible, to replace this rude teacher with a more gentle one: foresight (Frédéric Bastiat, 1848)

Predicting the future is notoriously difficult. Many great thinkers have tried, and spectacularly failed. For example, in 1895, Lord Kelvin, a Scottish mathematician and physicist is famous to have forthrightly stated to have ‘*not the smallest molecule of faith*’ in aerial flight beyond ballooning, just 8 years before the Wright brothers put together the first successful fixed wing aeroplane. Similarly, in a 1961 interview T.A.M. Craven the US Federal Communications commissioner of the time famously predicted: ‘*There is*

practically no chance communications space satellites will be used to provide better telephone, telegraph, television or radio service inside the United States’. Only a few years later, satellites were in space performing all the above services. So perhaps it is best to keep our heads below the parapet and not make predictions that, in retrospect, appear foolhardy?

In many fields of research, however, forecasting – prediction of future states based on past events, is essential. Forecasts allow us to alter our behaviours in response to likely realisations of future events in order to reduce costs or maximise benefits. In the environmental sciences, for example, weather forecasts, which have improved greatly in recent decades, provide huge overall benefit to society. In the longer term, climatological forecasts provide critical guidance to help steer our socioeconomic systems away from unsustainable and self-destructive pathways. There are still many dangers associated with ‘getting it wrong’ (as UK weather forecaster Michael Fish

*Corresponding author. E-mail: toliver@ceh.ac.uk

famously did in 1987 when he told people not to worry about a hurricane just hours before winds reaching 122 mph hit southern England). However, hiding away from making forecasts is often not an option, because this path leads to greater overall costs than making predictions that are occasionally wrong.

In ecological science, forecasting is also essential. It forms part of a set of tools, including horizon scanning (Sutherland *et al.*, 2008; Roy *et al.*, 2014) and risk assessments (Mace *et al.*, 2008; Thomas *et al.*, 2011), that enable us to anticipate future changes and respond appropriately. Reactive responses to new environmental impacts caused through changes in socioeconomic systems (e.g. adoption of new technologies), may often be too late to avert significant environmental damage.

For example, the pesticide DDT caused substantial losses to bird populations before it was finally banned (US Environmental Protection Agency, 1975; Pimentel, 2005). In Europe, agricultural subsidies, paid to farmers to increase food production and security, have led to increased loss of natural or semi-natural habitats, and have been a primary cause of European biodiversity decline (Van Swaay *et al.*, 2010; UK NEA, 2011; Inger *et al.*, 2014). The responses to mitigate these environmental impacts and others have mostly been reactive, in that they occurred only when damage had begun. In contrast, risk assessments based on experimental evidence of pesticide toxicity along with ecological modelling to predict potential impacts on species populations at larger spatial scales could have enabled proactive preventative measures to be taken.

In many cases, due to slow decision making and policy implementation, significant damage has been done before ameliorative actions are in place. The slow progress to develop co-operative global actions to halt climate change may turn out to be another such example, with potentially very large consequences for the environment and society (IPCC, 2014). In other cases, policy responses may be rapidly formulated based on hastily gathered evidence. In both situations there are strong benefits of early evidence gathering, ecological modelling and risk assessment to inform timely and evidence-based policy decisions. It should be recognised, however, that it will never be possible to foresee all ecological problems and some environmental management will have to be reactive.

In understanding the chain of events leading to environmental impacts, the 'DPSIR' (driver, pressure, state, impact, response) framework can be useful and is widely used (Fig. 1; European Environment Agency, 2007; United States Environmental Protection Agency, 2014). This framework illustrates the causal links between the ultimate drivers of environ-

mental degradation (e.g. population growth), the proximate pressures (e.g. food production) on the state of the environment (e.g. biodiversity) and their final impacts on humans (e.g. loss of well-being through degradation of ecosystem services that are underpinned by biodiversity). Societal responses may then be put in place to ameliorate these impacts. These may tackle the ultimate drivers (e.g. campaigns to educate on the environmental impacts of population growth) or proximate pressures (e.g. sustainable food production) or try to address the state of the environment directly without addressing drivers and pressures (e.g. improving the quality of semi-natural habitats that are known to support a high diversity of specialised species). However, the key problem with the DPSIR framework is that following it sequentially, as described above, amounts to reactive management practices that are often too late to avert significant environmental damage. It would be far better, as the quotation at the start of this article suggests, to be able to look forwards and predict possible impacts so that they can be averted. Thus, there is great need to forecast the impact of environmental drivers and pressures on the state of the environment.

EXTRAPOLATION OF TRENDS – THE SIMPLEST WAY TO FORECAST

By far the most straightforward way of predicting future states of the environment is to identify past trends in state over time and extrapolate these forwards using some kind of statistical model. For example the Global Biodiversity Outlook 4 report (GBO-4; Secretariat of the Convention on Biological Diversity, 2014) includes indicator-based extrapolations of recent and current trends to 2020. The report states 'The assessment of progress towards the Aichi Biodiversity Targets in GBO-4 is informed by recent trends in 55 biodiversity-related indicators and their statistical extrapolation to 2020'. Extrapolation can sometimes work well. For example, as

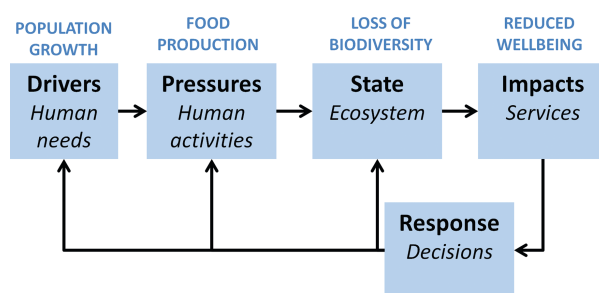


Figure 1. The 'DPSIR' framework with an example of drivers, pressures, state and impacts in capitalised font.

shown in Figure 2A and B for the butterfly species *Aphantopus hyperantus*, a linear trend fitted to a national population index from 1980 to 2000 adequately predicts abundances for the subsequent 13 years (mean absolute error = 0.076). In other cases, extrapolation does a poor job at predicting the future state of an environmental variable. This situation might be due to several reasons: (1) there is substantial error in our measurement of the system state; (2) our statistical model is inadequate (e.g. fitting a linear trend when there is significant curvature); (3) there is a high degree of short-term variability in the system state about some trend (sometimes called 'stochasticity'); or (4) the 'rules' that govern the system state change over time (i.e.

the drivers and pressures change). Examples of unsuccessful extrapolations are shown for the butterfly species *Euphydryas aurinia* (Fig. 2C, D; mean absolute error = 0.217) and *Hesperia comma* (Fig. 2E, F; mean absolute error = 0.421). In these cases, neither linear models, nor second or third order polynomials, fitted to 1980–2000 data are able to adequately predict population indices in the subsequent 13 years.

Why are extrapolations so poor for these species? The population indices are collated from a reasonably large number of sites (mean number of sites per year for each species \pm standard error: *A. hyperantus* = 324.7 ± 32.3 ; *E. aurinia* = 62.5 ± 5.24 ; *Hesperia comma* = 25.8 ± 1.9) and so any measurement

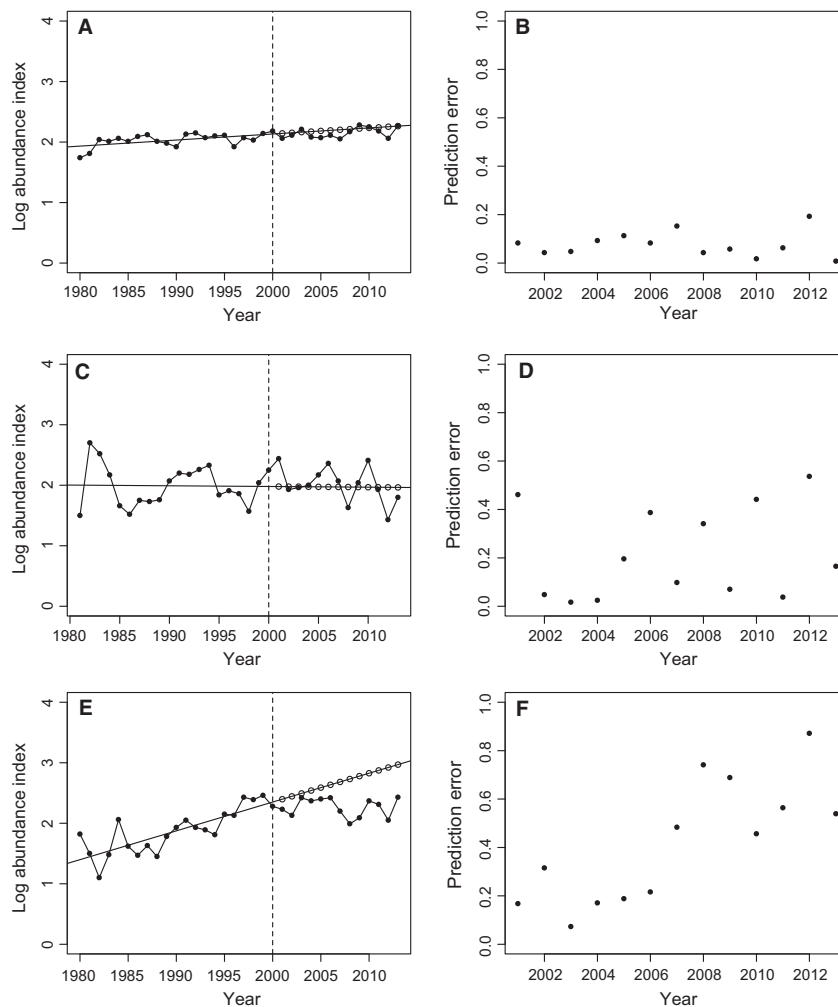


Figure 2. Extrapolations of population abundance for three butterfly species: the Ringlet *Aphantopus hyperantus* (A, B); Marsh Fritillary *Euphydryas aurinia* (C, D); and Silver Spotted Skipper *Hesperia comma* (E, F). Left hand panels show the UK national log collated index of abundance with a linear trend fitted to the data from 1980 to 2000 and used to predict abundance from 2001 onwards (open circles). The right hand panel shows the absolute difference between predicted and observed values.

errors should cancel each other out and have negligible effect. With regards to the statistical model, for each species three different models were compared (linear and second and third order polynomials) and in both cases the linear model was the best fit to the 1980–2000 data. In the case of *E. aurinia* (Fig. 2C, D), the interannual population variability is very large, and although the linear model is a better fit compared with models with curvature, the predicted values of the population index are often poor estimates. This factor may be problematic if accurate annual predictions are necessary; for example, if we were aiming to predict the abundance of a pest species in order to inform prophylactic pesticide application. It may be possible to predict some of the interannual variability in population abundances with models incorporating the factors that drive population dynamics (see next section), but there will often remain variation left over that we are unable to explain (e.g. resulting from demographic stochasticity). Despite this outcome, it is notable that errors in our predictions do not necessarily get markedly worse the further on in time a prediction is made (e.g. Fig. 2D; regression of absolute error by year: $F_{1,11} = 0.42$, $P = 0.53$). Therefore, if accurate annual predictions are not so important, but we are rather more interested in broad forecasts of future population trends, e.g. to allocate conservation funding appropriately, then large interannual population variability may not be such an issue (Roy *et al.*, 2001). The exception here would be for special cases in which stochasticity itself changes over time (e.g. changing environmentally induced stochasticity as populations move towards – or away – from the edge of their fundamental niche space as the climate changes; Oliver, Brereton & Roy, 2013; Oliver *et al.*, 2014).

Our third example (*Hesperia comma*; Fig. 2E, F), is also a poor forecast from extrapolation, but in this case the interannual population variability is relatively low (Bennie *et al.*, 2013). Instead, the linear trend, which is a good fit to the population indices between 1980 and 2000, is a poor fit to the data from 2001 onwards, with the predictions getting noticeably worse the further ahead we try to predict (Fig. 2F; regression of absolute error by year: $F_{1,11} = 18.2$, $P = 0.001$). The population trajectory has changed direction, presumably because density-dependent processes are beginning to operate or because the primary drivers that affect populations (e.g. climate, habitat quality, habitat extent) have changed over the duration of the monitoring period. This problem is critical for extrapolation methods, because by using data only on the system state we cannot account for changes in drivers. Instead, they are assumed to be constant; an assumption that is

very often contravened. For example, Mason *et al.* (2015) consider rates of distribution change (northern range margin shift) in four different animal groups over two time intervals (from 1970 to 2010) and find that rates of change in the first time interval are poor predictors of rates of change in the subsequent time interval. What are these changes in drivers that underlie species responses?

The UK National Ecosystem Assessment (UK NEA, 2011) was the first comprehensive review of drivers of change in biodiversity and other ecosystem services. It assessed the historic impact of drivers, but also their expected future impact. It is notable that the magnitude of drivers often changes over time. For example, habitat loss and pollution have been the primary causes of biodiversity loss in the UK over the last century, but the impacts of these drivers are expected to lessen, with climate change and invasive species becoming the major new drivers of change (UK NEA, 2011). Similar patterns are likely to be occurring across other heavily modified temperate landscapes. For example, Carvalheiro *et al.* (2013) suggest that pollinator declines in a number of northwest European countries may have slowed, probably due to a peak in the conversion of land use to intensive agriculture. In addition to changes in the magnitude of existing drivers, new drivers may emerge with the advent of new technologies (August *et al.*, 2015). For example, a recent horizon scanning exercise by Sutherland *et al.* (2008) identified nanotechnology and geoengineering as fields with a large potential to impact biodiversity. Finally, drivers of change also interact in their environmental impacts, leading to non-linearity in responses (Brook, Sodhi & Bradshaw, 2008). For example, climate and land-use change can interact on biodiversity through a wide range of mechanisms affecting processes from demography to metapopulation structure and community interactions (Oliver & Morecroft, 2014). All these changes in drivers mean that simple extrapolations of system state are often a poor method of forecasting, especially over longer timescales. Instead, statistical models are needed that can incorporate the impact of drivers and how these may change in the future.

FORECASTING CHANGES IN DRIVERS AND PRESSURES

In order to incorporate the impact of pressures on the state of environmental systems we need to understand their functional relationships (e.g. what is the relationship between weather variables and a species' population size), and also anticipate how pressures are likely to change in the future (e.g.

what will future weather be like under climate change). Effects may be direct (e.g. weather impacts on demographic rates) or indirect (e.g. mediated through impacts on other species that interact with the focal species). The former task of understanding causal relationships can be achieved through experimentation (e.g. Tilman *et al.*, 1994) or through observation of ‘natural experiments’ (i.e. using long-term monitoring of system state and relating this to naturally occurring changes in drivers and pressures; Baker *et al.*, 2012; Eglington & Pearce-Higgins, 2012; Roy *et al.*, 2001). Both these methods are costly and time consuming, but the experimental designs (e.g. split-plot field experiments, long-term monitoring networks) and statistical analysis techniques needed (e.g. multivariate regressions, hierarchical mixed modelling and structural equation modelling) are well versed in the ecological sciences. In contrast, the latter task of anticipating future changes in drivers and pressures is more difficult and less practiced. One technique would be to use extrapolative techniques to predict how drivers and pressures may change based on past temporal trends. For example the GBO-4 report (Secretariat of the Convention on Biological Diversity, 2014) uses extrapolations of trends in human population size, gross domestic product, intensity of resource use, agricultural subsidies and surplus nitrogen in the environment. However, the problems with extrapolation of the state of the environment outlined above also hold true for extrapolation of pressures affecting system state – the pressures themselves are affected by multiple other factors (drivers) that may have non-linear trends over time. These more proximate drivers are themselves affected by other drivers, and so on. For example, nitrogen deposition is affected by both the cost of petrochemical fertilisers and price of crops in the world market, these factors themselves are affected by other more ultimate drivers such as population growth and development of alternative resource extraction and food production technologies. Suddenly we are faced with an enormous task: to forecast how a given pressure might change we need to understand the whole chain of causality affecting that pressure (the ‘infinite regress of drivers’ dilemma). Much ecological science, and indeed science in general, has tended to be reductionist in its approach, focusing on specific causal relationships, but more systematic ways of thinking are evident in ancient eastern world views (e.g. the Buddhist concept of ‘Pratītyasamutpāda’) that strongly emphasises the interdependency of entities and multiple causal linkages between them, and also feature in ‘systems ecology’ approaches (Odum, 1983; Schellnhuber, 1999; Evans *et al.*, 2013). Yet, there are clear practical limitations to understanding

changes to specific pressures by tracing causal links across the entire socio-ecological-economic system; effectively, this relies on a statistical model of the entire world!

At this point, one might be tempted to throw in the towel and give up trying to forecast future environmental states. However, returning to our original reason for attempting forecasting – that without it we rely on reactive management that is often too late to avert significant environmental damage – we are reminded that forecasting, although difficult, is very necessary. What is needed is a practical way forward that still remains as rigorous as possible. Pragmatic approaches to the problem so far have tended to either focus on one specific chain of causality affecting the environmental state (e.g. how climate change will affect local temperatures and how these will affect species populations; Thomas *et al.*, 2004), or to adopt a ‘storyline’ scenario approach where various possible socioeconomic scenarios are described and then a deliberative approach used to translate how changes in more ultimate drivers (such as population size, climate, new technologies) will impact on proximate pressures that affect environmental states. This latter approach is adopted by the Millennium Ecosystem Assessment (2005) and also in UK National Ecosystem Assessment (UK NEA, 2011) in order to explore how different socioeconomic scenarios might affect the state of the environment and the ecosystem services it provides. Both approaches have advantages but also some key problems.

The first approach that focuses on a single ‘chain’ of causality can afford more detailed quantitative analysis, but because this approach is more reductionist it ignores the importance of other interacting drivers and pressures on the system state. This may be warranted where there is good evidence for an overwhelmingly strong influence of one driver, which explains a large proportion of variance in the system state. For example, Figure 3 illustrates predictive correlative models for the abundance of three butterfly species, whose population dynamics are driven by annual weather and density dependence to varying extents. The first species, *Melanargia galathea* (A and B), shows a reasonably good fit between predicted and observed abundance in the model calibration period ($R^2 = 0.68$ with density dependence effects incorporated in models, $R^2 = 0.53$ without), and also in the model validation (mean absolute error [MAE] = 0.14–0.15) and the ‘forecasting’ periods (MAE = 0.10–0.11). In contrast, for the species *Celastrina argiolus* (C and D) even though goodness-of-fit for the model incorporating density dependence is greater than the one without ($R^2 = 0.72$ compared with 0.54) and there is little apparent difference in model fit in the validation period (MAE = 0.21

compared with 0.17), it turns out that the simpler model without density dependence performs slightly better in the forecasting period (MAE = 0.29 compared with 0.4). Interestingly, this species is one in which population dynamics are anecdotally thought to be driven by parasitoids, causing what looks like 6- to 8-year cycles (Revels, 2006). In fact, Figure 3D shows that a model with just three weather variables (spring rainfall in the current year, and autumn temperatures in the previous 2 years) can predict the population dynamics reasonably well. In other cases, it may be much less easy to predict abundances using weather variables. This can be despite an apparent good fit of models in the calibration period. For example, the species *Lasiommata megara* (Fig. 3E, F) has a model with a high goodness-of-fit between predicted and observed abundance in the

model calibration period ($R^2 = 0.83$ for the model also incorporating density dependence effects; Fig. 3E), yet forecasts are very poor (actually predicting local extinction by 2004). In this case, it would be clear from validation that the model without density dependence is actually a better predictor (MAE = 0.39 compared with 3.96) despite the markedly lower goodness-of-fit in the calibration period ($R^2 = 0.13$). This outcome might be because density-dependent processes are changing over time (e.g. due to changes in the parasites or pathogens that drive density dependence). Importantly, in 'free running' predictions (in which abundance is sequentially estimated as a function of abundance in the previous year) any errors in modelled density dependence relationships quickly accumulate, leading to very large prediction errors in just a few years.

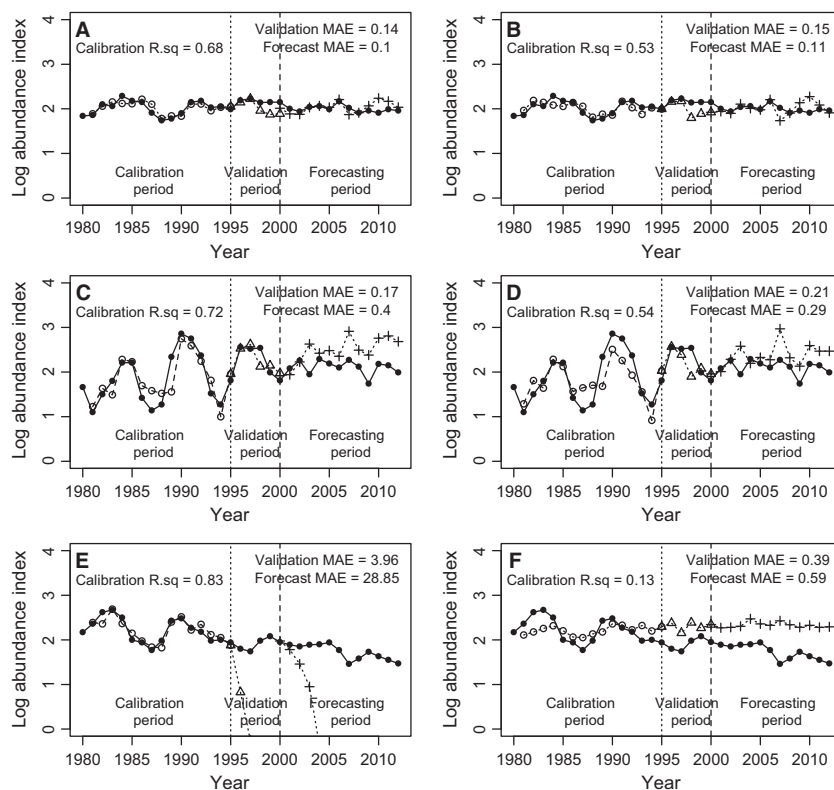


Figure 3. Observed (filled circles) and predicted annual abundance for three butterfly species: *Melanargia galathea* (A, B); *Celastrina argiolus* (C, D); and *Lasiommata megara* (E, F). Left hand panels show multiple regression models fitted to log abundance indices in a model calibration period from 1980 to 1995. The models include the three most important weather variables for each species (G. Palmer, J.K. Hill, T.M. Brereton, D.R. Brooks, J.W. Chapman, R. Fox, T.H. Oliver, C.D. Thomas submitted), and the previous years' density as explanatory variables. Each model is then used for hindcasting abundances in the calibration period using the observed density the previous year along with the observed weather variables (open circles). The goodness-of-fit (R^2) of the model is shown in the top left corner of each panel. Also shown are free running abundance predictions for a validation period from 1995 to 2000 (open triangles) and a 'forecasting' period from 2000 to 2012 (crosses). The mean absolute error of predictions is given in the top right hand corner of each panel. The right hand panels (B, D, F) are similar, but the models are fitted to only the three most important weather variables for the species without a density-dependence term.

These examples highlight the inherent difficulties of ecological forecasting. A good fit of models in calibration periods is not a sufficient indicator of good predictive ability, and only after successful validation should models be used for prediction. In practice, prediction errors would be larger than shown here because true forecasts would be based on estimates of weather from downscaled climate models, rather than the observed weather variables used here. Even with observed weather data and even though ectothermic butterflies tend to be highly responsive to weather, these correlative statistical models based on large datasets still only had limited predictive ability. Relationships with density and weather could change over time (e.g. due to changes in the factors causing density dependence, such as the aggregation of natural enemies, and due to evolutionary adaptation of populations to changes in climate). In addition, other drivers beyond climate are likely to be important in driving population changes (e.g. habitat quality).

The second ‘storyline’ scenario approach to predicting the impacts of environmental change has the advantage of considering many different interacting drivers and pressures. The scenarios generated are not necessarily taken as predictions of the future but, rather, delineate a wide range of possible outcomes in the parameter space of multiple drivers and pressures. The major disadvantage of this approach, however, is that the large number of causal relationships between multiple drivers and environmental state, including interaction effects, means that predictive models are nearly impossible to parameterise empirically; instead, these rely on expert opinion, or a combination of both such as Bayesian belief networks (Haines-Young *et al.*, 2011). The other key problem is that the models are so complex that any resulting differences in environmental state between scenarios are difficult to attribute back to specific drivers. For example, the UK National Ecosystem Assessment includes various storyline scenarios such as ‘Nature at Work’ and the ‘National Security’ that differ greatly in predicted ecosystem services (Bateman *et al.*, 2011; Haines-Young *et al.*, 2011). However, it is impossible to say whether this is due to differences in protected area coverage, the extent of several different cover types, or due to varying impacts of climate change, which all differ between the scenarios.

For successful ecological forecasting, it is necessary to find a practical compromise between the two extremes of quantitative reductionism (which ignores multiple interacting drivers) and a fully systemic approach (which is too complex to parameterise and to attribute final changes in environment state). Below we suggest a practical middle ground

that may potentially reconcile these extreme approaches.

KEY DRIVER AND RESPONSE –TEST SCENARIOS

One possible middle-ground solution for ecological forecasting lies in identifying the minimum number of key drivers or pressures that is demonstrated (through empirical analysis or expert opinion) to strongly influence an environmental state variable over the time frame of interest. This solution avoids the excessive complexity of considering all drivers in models, but also extends beyond the simplistic view of considering change in one driver in isolation. The suite of drivers considered may already be operating or they may be potential new drivers. In addition, interactions between drivers should also be considered when relevant. In order to make our ecological forecast models of applied use, we should also consider alternative possible ‘response’ options (i.e. management solutions). For example, in considering the possible impact of climate change on populations of a species we might consider a range of land-use scenarios that, through the existence of land-use climate interactions, potentially allows the amelioration of climate change impacts (i.e. ‘adaptation’ measures).

WHICH STATISTICAL METHOD TO USE?

Thus far, we have broadly considered ecological forecasting methods in terms of the merits of trend extrapolation techniques versus multivariate models that incorporate changes in drivers and pressures. The latter method is recommended with a pragmatic approach to selecting key drivers and incorporating response options. However, we have so far discussed specific statistical methods only fleetingly. A wide range of statistical methods is relevant to this problem, ranging from simple correlative models to highly complex individual-based ‘process’ models. Table 1 gives examples of a range of methods that spans a continuum of complexity and input data requirements (also see Sutherland, 2006). Ecological modellers often specialise on some part of this model continuum. For example, fitting correlative models across many species (macroecology) versus developing detailed individual-based models for specific species (process-based modelling). Neither approach is right or wrong and each one has its advantages and disadvantages. Detailed process-based models are often highly complex, aiming to incorporate many biological processes (e.g. demographic parameters

Table 1. Different statistical methods for ecological forecasting

Forecasting method	Description	Example
Extrapolation	Descriptive statistical model of trend in system state variable	Predicting species geographic range margins from past rates of change
Simple correlative models	Statistical relationship between driver or pressure variable(s) and system state variable	Predicting species distribution from relationships between occurrence and climate variables
Phenomenological models*	Statistical relationship between driver or pressure and intermediate demographic processes that combine to determine state	Predicting species abundance from climate impacts on population growth, mortality and dispersal
Mechanistic model*	Relationship between driver or pressure and intermediate demographic processes based on prior biological understanding	Predicting species abundance under climate changes based on physiological relationships between development rates and lethal temperatures
Individual-based models	Behavioural rules used to model individual decisions, often in combination with phenomenological components relating to demography	Predicting impacts of climate change on individual movements and how this scales up to species range margin shifts

*Note that this dichotomy is not strict and some models combine both phenomenological and mechanistic components.

that vary in different environments, interspecific interactions and evolutionary processes). As such, they are more biologically realistic representations with the potential for more accurate predictions. However, the cost of this complexity is that many more parameters need to be estimated, requiring significantly more data to calibrate models. When one considers the total numbers of species that one could potentially be interested in for conservation ecology (e.g. c. 70 000 in the UK, which is a relatively species-poor country; UK Species Inventory, 2014; Gurney, 2015), then it becomes clear that such data-hungry models are impractical, unless they can be shown to produce general responses that are representative of many other species. At the other end of the spectrum, extrapolative techniques or simple correlative models, which consider species responses to a single driver, may be too simplistic and ignore key interactions between drivers. This lack of mechanistic understanding behind species responses can mean that predictions are inaccurate if drivers change in non-linear ways over time. To identify a practical way forward, again, the theory of the ‘middle way’ may help up to reconcile these extremes. Methods are needed that balance the complexity needed to make robust predictions with the feasibility of model parameterisation given data availability. The three methods in the middle rows of Table 1 are most likely to achieve this balance and produce reliable forecasts of environmental change for many species in order to inform conservation responses. These include phenomenological models and mechanistic models. In practice, successful models may contain a combination of both the above approaches, with well

known biological processes specified by mechanistic relationships but with flexibility for unknown relationships to be estimated from the observed data (Dormann *et al.*, 2012).

RIGOUR IN PREDICTIVE MODELLING

In addition to selecting the most appropriate statistical modelling framework, the modelling approach must be as rigorous as possible, in order to ensure accurate predictions. This situation is especially important if our forecasts are used as evidence to implement prophylactic management to avert environmental damage. Such management options may have substantial costs and therefore need to be well evidenced. For example, setting aside semi-natural habitats to maintain pollinating insects under climate change has costs in terms of reduced land for cropping and so strong evidence is needed to convince stakeholders of the best land management solution.

To select the most appropriate statistical model, the goodness-of-fit to historic data of alternative models is usually assessed. But this criterion alone can lead to over-parameterised models that are poor at predicting future environmental states. Instead, model validation is necessary using data independent of that used for model fitting. For example, Figure 3 shows how 20 years of historic monitoring data can be split into 15 years for model fitting (of abundance changes in relation to weather variables) and five for model testing. This approach can prevent overfitting and allow better predictions of subsequent

abundance (demonstrated here by hindcasting to the most recent 13 years of data). In the absence of time series data, space-for-time substitutions are the next best option to validate models; for example, using the model fitted to data in one area to predict environmental state in a different area. There can be problems with this approach, however, as a number of different correlated variables may change across spatial gradients (White & Kerr, 2006; Isaac *et al.*, 2010).

Despite the importance of validation, yet because of its difficulty, several current modelling frameworks are being used to predict future environmental states with limited validation of the models. This approach is especially evident in the recently emerging field of ecosystem service modelling (e.g. Nelson *et al.*, 2009, 2010; UK NEA, 2011; Bateman *et al.*, 2013). There is clearly a danger in implementing management options with limited evidence. Under such circumstances an adaptive management approach is highly appropriate, in which the management actions predicted as most suitable are taken but with regular monitoring to assess their effectiveness. Past environmental policies have often tended to be inflexible, however, leading to 'lock-in' to a set of options with little consideration of adaptive management. For example, agri-environment schemes put in place in the UK to prevent biodiversity declines in agricultural landscapes were based on synthesis of the evidence base for the effectiveness of different management options. Over £400 M per year is spent on these schemes in England (Natural England, 2009), yet the budget to monitor the effectiveness of these schemes (and validate the predictions of their effectiveness) is a very small percentage of this (< 1%). With this and many other environmental policies, when seen in the context of ecological forecasting and its validation, there is a strong argument for rebalancing spending on action versus monitoring and analysis.

DATA FOR PREDICTIVE MODELS AND THE IMPORTANCE OF BIOLOGICAL RECORDING

As discussed above, a wide range of environmental drivers can impact species. Models of these causal relationships (and the potential interactions between drivers) will necessarily have several estimated parameters, even when reduced to the subset of drivers with the largest impacts. Therefore, substantial datasets on species populations and measured values of drivers are needed for model calibration and validation. These data may come from mesocosm or field experiments, or 'natural experiments' comprising the monitoring of natural populations over broad

environmental gradients. Mesocosm experiments consider responses to a limited range of manipulated variables under controlled conditions. They are useful for testing theory and stimulating further research (Benton *et al.*, 2007), although the ability of these systems to produce species responses similar to those in the real world is questionable (Carpenter, 1996). Field experiments comprise a selected range of treatments to consider the effects of different drivers but at a larger scale and in more realistic settings subject to 'noise' from other unmeasured environmental drivers (Carpenter, 1998). Experimental manipulation is the ideal way to test ecological theories (including those pertaining to relationships between drivers and population responses), but field experiments are very time consuming and expensive. In addition, even in large-scale experiments, the limited spatial and temporal scale means that patterns operating at large scales may be missed (e.g. Wiens, Rotenberry & Van Horne, 1986). The alternative to designed experiments is to exploit natural environmental gradients and large-scale perturbations (Carpenter, 1990). This approach requires monitoring of species responses, ideally with high levels of spatial and temporal replication that are necessary to maintain statistical power in the face of combined variability across wide range of environmental variables. Examples of such large-scale monitoring include the collection of biological records (georeferenced records of species presences), such as those held by the Biological Records Centre (Pocock *et al.*, 2015), the broad utility of which is demonstrated by the articles within this issue (Chapman *et al.*, 2015; Gillingham *et al.*, 2015; Mason *et al.*, 2015; Powney & Isaac, 2015; Purse & Golding, 2015; Roy, 2015; Roy *et al.*, 2015; Sutherland, Roy & Amano, 2015). Abundance data such as that represented in species monitoring schemes (e.g. the UK Butterfly Monitoring Scheme data used in this paper), are even more useful in allowing forecasts of abundance rather than just species presence, although there is evidence that distribution data can predict abundance to a limited degree (Elmendorf & Moore, 2008; VanDerWal *et al.*, 2009; Oliver *et al.*, 2012a,b). Monitoring data provide a crucial resource for ecological forecasting because of the large spatial and temporal extent that they can cover. This resource is facilitated by the use of trained volunteers that help to reduce the total costs of monitoring schemes. Experimental approaches, although much better for well controlled tests of theory and testing management techniques, are often too limited in their spatial coverage to adequately inform ecological forecasts, at least beyond the location of the experiment. For example, species responses to weather conditions (a major driver of population variation across most species) can be

assessed using data from national monitoring schemes (Roy *et al.*, 2001; Wallis De Vries, Baxter & Van Vliet, 2011; Eglinton & Pearce-Higgins, 2012). In addition, because weather effects are modified by topography and habitat type not just at the local scale but also by the structural composition of surrounding landscapes (Oliver *et al.*, 2010, 2012a, 2013), then it is necessary to empirically model these effects in order to make general forecasts beyond the responses of a single site. Fortunately, some countries have well established species monitoring schemes (e.g. the long history of natural history recording in Britain), and protocols for a number of currently unstudied species groups and new monitoring schemes in other countries are currently under development. International initiatives are also at work to synthesise monitoring data (e.g. the global Biodiversity Observation Network, GEOBON).

CONCLUSION

Overall, this review has highlighted the necessity of ecological forecasting in order to avert environmental damages that would occur under a solely reactive management approach. Extrapolating from historic environmental states will often be unsuccessful because drivers and pressures themselves change in non-linear ways and interact in their impact. New pressures on systems, such as those from emerging technologies, may also arise. The complexity of causal pathways between drivers, pressures and environmental state is effectively endless, so a pragmatic approach is needed to focus on a subset of pathways that have the greatest impact on system state. Additionally, incorporating possible response options in our models will allow us to assess the effectiveness of potential solutions to environmental damage. In terms of modelling techniques, a compromise between complexity (and potential biological realism) and feasibility given data availability will be necessary, but there is also a key importance in validating models to ensure that forecasts can be made with reasonable confidence. Although ecological forecasting is very difficult, it is also highly necessary. Much ecological science has tended to focus on understanding current and historic patterns and trends. However, there is a clear need to step out of our comfort zones and develop ecological forecasts in order to inform environmental management effectively.

ACKNOWLEDGEMENTS

We are grateful to the UK Butterfly Monitoring Scheme (UKBMS) volunteers and scheme co-ordinators

for providing the data used in the forecasting examples in this paper. The UKBMS is funded funding by a multi-agency consortium led by the Department for Environmental Food and Rural Affairs, and including the Natural Resources Wales, Natural England, Forestry Commission, and Scottish Natural Heritage. The Biological Record Centre, of which the UKBMS is part, receives support from the Joint Nature Conservation Committee and the Natural Environment Research Council (via National Capability funding to the Centre for Ecology & Hydrology, project NEC04932). We thank Mike Morecroft and Chris Preston for helpful comments that improved this manuscript.

REFERENCES

- August T, Harvey M, Lightfoot P, Kilbey D, Papadopoulos T, Jepson P. 2015. Emerging technologies for biological recording. *Biological Journal of the Linnean Society* **115**: 731–749.
- Baker DJ, Freeman SN, Grice PV, Siriwardena GM. 2012. Landscape-scale responses of birds to agri-environment management: a test of the English Environmental Stewardship scheme. *Journal of Applied Ecology* **49**: 871–882.
- Bateman IJ, Abson D, Andrews B, Crowe A, Darnell A, Dugdale S, Fezzi C, Foden J, Haines-Young R, Hulme M, Munday P, Pascual U, Paterson J, Perino G, Sen A, Siriwardena G, Termansen M. 2011. Chapter 26: valuing changes in ecosystem services: scenario analyses. In: *The UK national ecosystem assessment: technical report*. Cambridge: UNEP-WCMC. Available at: <http://uknea.unep-wcmc.org/LinkClick.aspx?fileticket=m%2BvhAV3c9uk%3D&tabid=82>.
- Bateman IJ, Harwood AR, Mace GM, Watson RT, Abson DJ, Andrews B, Binner A, Crowe A, Day BH, Dugdale S, Fezzi C, Foden J, Hadley D, Haines-Young R, Hulme M, Kontoleon A, Lovett AA, Munday P, Pascual U, Paterson J, Perino G, Sen A, Siriwardena G, van Soest D, Termansen M. 2013. Bringing ecosystem services into economic decision-making: land use in the United Kingdom. *Science* **341**: 45–50.
- Bennie J, Hodgson JA, Lawson CR, Holloway CTR, Roy DB, Brereton T, Thomas CD, Wilson RJ. 2013. Range expansion through fragmented landscapes under a variable climate. *Ecology Letters* **16**: 921–929.
- Benton TG, Solan M, Travis JMJ, Sait SM. 2007. Microcosm experiments can inform global ecological problems. *Trends in Ecology & Evolution* **22**: 516–521.
- Brook BW, Sodhi NS, Bradshaw CJA. 2008. Synergies among extinction drivers under global change. *Trends in Ecology and Evolution* **23**: 453–460.
- Carpenter SR. 1990. Large-scale perturbations: opportunities for innovation. *Ecology* **71**: 2038–2043.
- Carpenter SR. 1996. Microcosm experiments have limited relevance for community and ecosystem ecology. *Ecology* **77**: 677–680.

- Carpenter SR. 1998.** The need for large-scale experiments to assess and predict the response of ecosystems to perturbation. In: Pace M, Groffman P, eds. *Successes, limitations, and frontiers in ecosystem science*. New York: Springer, 287–312.
- Carvalho LG, Kunin WE, Keil P, Aguirre-Gutiérrez J, Ellis WN, Fox R, Groom Q, Hennekens S, Van Landuyt W, Maes D, Van de Meutter F, Michez D, Rasmont P, Ode B, Potts SG, Reemer M, Roberts SPM, Schaminée J, WallisDeVries MF, Biesmeijer JC. 2013.** Species richness declines and biotic homogenisation have slowed down for NW-European pollinators and plants. *Ecology Letters* **16**: 870–878.
- Chapman D, Bell S, Helfer S, Roy DB. 2015.** Unbiased inference of plant flowering phenology from biological recording data. *Biological Journal of the Linnean Society* **115**: 543–554.
- Dormann CF, Schymanski SJ, Cabral J, Chuine I, Graham C, Hartig F, Kearney M, Morin X, Römermann C, Schröder B, Singer A. 2012.** Correlation and process in species distribution models: bridging a dichotomy. *Journal of Biogeography* **39**: 2119–2131.
- Eglinton S, Pearce-Higgins JW. 2012.** Disentangling the relative importance of changes in climate and land-use intensity in driving recent bird population trends. *PLoS ONE* **7**: e30407.
- Elmendorf SC, Moore KA. 2008.** Use of community-composition data to predict the fecundity and abundance of species. *Conservation Biology* **22**: 1523–1532.
- European Environment Agency. 2007.** The DPSIR framework used by the EEA. Available at: http://ia2dec.eea.europa.eu/knowledge_base/Frameworks/doc101182.
- Evans MR, Bithell M, Cornell SJ, Dall SRX, Diaz S, Emmott S, Ernande B, Grimm V, Hodgson DJ, Lewis SL, Mace GM, Morecroft M, Moustakas A, Murphy E, Newbold T, Norris KJ, Petchey O, Smith M, Travis JMJ, Benton TG. 2013.** Predictive systems ecology. *Proceedings of the Royal Society B: Biological Sciences* **280**.
- Gillingham PK, Bradbury RB, Roy DB, Anderson BJ, Baxter JM, Bourn NAD, Crick HQP, Findon RA, Fox R, Franco A, Hill JK, Hodgson JA, Holt AR, Morecroft MD, O'Hanlon NJ, Oliver TH, Pearce-Higgins JW, Procter DA, Thomas JA, Walker KJ, Walmsley CA, Wilson RJ, Thomas CD. 2015.** The effectiveness of protected areas in the conservation of species with changing geographical ranges. *Biological Journal of the Linnean Society* **115**: 707–717.
- Gurney M. 2015.** Gains and losses: recent colonisations and extinctions in Britain. *Biological Journal of the Linnean Society* **115**: 573–585.
- Haines-Young R, Paterson J, Potschin M, Wilson A, Kass G. 2011.** Chapter 25: The UK NEA Scenarios: Development of Storylines and Analysis of Outcomes. In: *The UK national ecosystem assessment: synthesis of the key findings*. Cambridge: UNEP-WCMC. Available at: <http://uknea.unep-wcmc.org/LinkClick.aspx?fileticket=m%2BvhAV3c9uk%3D&tabid=82>
- Inger R, Gregory R, Duffy JP, Stott I, Vorišek P, Gaston KJ. 2014.** Common European birds are declining rapidly while less abundant species' numbers are rising. *Ecology Letters* **18**: 28–36.
- IPCC. 2014.** Working Group II Contribution to the IPCC Fifth Assessment Report. Climate Change 2014: Impacts, Adaptation, and Vulnerability.
- Isaac NJB, Girardello M, Brereton T, Roy D. 2010.** Butterfly abundance in a warming climate: patterns in space and time are not congruent. *Journal of Insect Conservation* **15**: 233–240.
- Mace GM, Collar NJ, Gaston KJ, Hilton-Taylor C, Akçakaya HR, Leader-Williams N, Milner-Gulland EJ, Stuart SN. 2008.** Quantification of extinction risk: IUCN's system for classifying threatened species. *Conservation Biology* **22**: 1424–1442.
- Maes D, Isaac NJB, Harrower CA, Collen B, van Strien AJ, Roy DB. 2015.** The use of opportunistic data for IUCN Red List assessments. *Biological Journal of the Linnean Society* **115**: 690–706.
- Mason SC, Palmer G, Fox R, Gillings S, Hill JK, Thomas CD, Oliver TH. 2015.** Geographical range margins of many taxonomic groups continue to shift polewards. *Biological Journal of the Linnean Society* **115**: 586–597.
- Millennium Ecosystem Assessment. 2005.** *Ecosystems and Human Well-being: Opportunities and Challenges for Business and Industry*. Washington, DC: World Resources Institute.
- Natural England. 2009.** Agri-environment schemes in England 2009: a review of results and effectiveness: Natural England Report (NE194).
- Nelson E, Mendoza G, Regetz J, Polasky S, Tallis H, Cameron D, Chan KMA, Daily GC, Goldstein J, Kareiva PM, Lonsdorf E, Naidoo R, Ricketts TH, Shaw M. 2009.** Modeling multiple ecosystem services, biodiversity conservation, commodity production, and tradeoffs at landscape scales. *Frontiers in Ecology and The Environment* **7**: 4–11.
- Nelson E, Sander H, Hawthorne P, Conte M, Ennaanay D, Wolny S, Manson S, Polasky S. 2010.** Projecting global land-use change and its effect on ecosystem service provision and biodiversity with simple models. *PLoS ONE* **5**: e14327.
- Odum HT. 1983.** *Systems Ecology: An Introduction*. Wiley-Interscience.
- Oliver T, Roy DB, Hill JK, Brereton T, Thomas CD. 2010.** Heterogeneous landscapes promote population stability. *Ecology Letters* **13**: 473–484.
- Oliver TH, Roy DB, Brereton T, Thomas JA. 2012a.** Reduced variability in range-edge butterfly populations over three decades of climate warming. *Global Change Biology* **18**: 1531–1539.
- Oliver TH, Gillings SG, Girardello M, Rapacciuolo G, Brereton T, Siriwardena GM, Roy DB, Pywell RF, Fuller RJ. 2012b.** Population density but not stability can be predicted from species distribution models. *Journal of Applied Ecology* **49**: 581–590.
- Oliver TH, Brereton T, Roy DB. 2013.** Population resilience to an extreme drought is influenced by habitat area and fragmentation in the local landscape. *Ecography* **36**: 579–586.
- Oliver TH, Morecroft MD. 2014.** Interactions between climate change and land use change on biodiversity: attribu-

- tion problems, risks, and opportunities. *Wiley Interdisciplinary Reviews: Climate Change* **5**: 317–335.
- Oliver TH, Stefanescu C, Paramo F, Brereton T, Roy DB. 2014.** Latitudinal gradients in butterfly population variability are influenced by landscape heterogeneity. *Ecography* **37**: 863–871.
- Pocock MJO, Roy HE, Preston CD, Roy DB. 2015.** The biological records centre: a pioneer of citizen science. *Biological Journal of the Linnean Society* **115**: 475–493.
- Powney GD, Isaac NJB. 2015.** Beyond maps: a review of the applications of biological records. *Biological Journal of the Linnean Society* **115**: 532–542.
- Purse B, Golding N. 2015.** Tracking the spread and impacts of diseases with biological records and distribution modelling. *Biological Journal of the Linnean Society* **115**: 664–677.
- Pimentel D. 2005.** Environmental and economic costs of the application of pesticides primarily in the United States. *Environment, Development and Sustainability* **7**: 229–252.
- Revels R. 2006.** More on the rise and fall of the Holly blue. *British Wildlife* **17**: 419–424.
- Roy DB, Rothery P, Moss D, Pollard E, Thomas JA. 2001.** Butterfly numbers and weather: predicting historical trends in abundance and the future effects of climate change. *Journal of Animal Ecology* **70**: 201–217.
- Roy HE, Peyton J, Aldridge DC, Bantock T, Blackburn TM, Britton R, Clark P, Cook E, Dehnen-Schmutz K, Dines T, Dobson M, Edwards F, Harrower C, Harvey MC, Minchin D, Noble DG, Parrott D, Pocock MJO, Preston CD, Roy S, Salisbury A, Schönrogge K, Sewell J, Shaw RH, Stebbing P, Stewart AJA, Walker KJ. 2014.** Horizon scanning for invasive alien species with the potential to threaten biodiversity in Great Britain. *Global Change Biology* **20**: 3859–3871.
- Roy HE, Rorke SL, Beckmann B, Botham MS, Brown PMJ, Noble D, Sewell J, Walker KJ. 2015.** The contribution of volunteer recorders to our understanding of biological invasions. *Biological Journal of the Linnean Society* **115**: 678–689.
- Schellnhuber HJ. 1999.** ‘Earth system’ analysis and the second Copernican revolution. *Nature* **402**: C19–C23.
- Secretariat of the Convention on Biological Diversity. 2014.** *Global Biodiversity Outlook 4*: Montréal.
- Sutherland WJ, Roy DB, Amano T. 2015.** An agenda for the future of biological recording for ecological monitoring and citizen science. *Biological Journal of the Linnean Society* **115**: 779–784.
- Sutherland WJ. 2006.** Predicting the ecological consequences of environmental change: a review of the methods. *Journal of Applied Ecology* **43**: 599–616.
- Sutherland WJ, Bailey MJ, Bainbridge IP, Brereton T, Dick JTA, Drewitt J, Dulvy NK, Dusic NR, Freckleton RP, Gaston KJ, Gilder PM, Green RE, Heathwaite AL, Johnson SM, Macdonald DW, Mitchell R, Osborn D, Owen RP, Pretty J, Prior SV, Prosser H, Pullin AS, Rose P, Stott A, Tew T, Thomas CD, Thompson DBA, Vickery JA, Walker M, Walmsley C, Warrington S, Watkinson AR, Williams RJ, Woodroffe R, Woodroof HJ. 2008.** Future novel threats and opportunities facing UK biodiversity identified by horizon scanning. *Journal of Applied Ecology* **45**: 821–833.
- Thomas CD, Cameron A, Green RE, Bakkenes M, Beaumont LJ, Collingham YC, Erasmus BFN, Ferreira de Siqueira M, Grainger A, Hannah L, Hughes L, Huntley B, van Jaarsveld AS, Midgley GF, Miles L, Ortega-Huerta MA, Peterson AT, Phillips OL, Williams SE. 2004.** Extinction risk from climate change. *Nature* **427**: 145–148.
- Thomas CD, Hill JK, Anderson BJ, Bailey S, Beale CM, Bradbury RB, Bullman CR, Crick HPQ, Eigenbrod F, Griffiths H, Kunin WE, Oliver TH, Walmsley CA, Watts K, Worsfold NT, Yardley T. 2011.** A framework for assessing threats and benefits to species responding to climate change. *Methods in Ecology and Evolution* **2**: 125–142.
- Tilman D, Dodd M, Silvertown J, Poulton P, Johnston A, Crawley M. 1994.** The park grass experiment—insights from the most long-term ecological study. In: Leigh RA, Johnston AE, eds. *Long-term Experiments in Agricultural and Ecological Science*. Wallingford, Oxon. Wallingford, Oxon, UK: CAB International, 287–303.
- UK NEA. 2011.** *The UK national ecosystem assessment: synthesis of the key findings*. Cambridge: UNEP-WCMC.
- UK Species Inventory. 2014.** *Natural History Museum*. Available at: <http://www.nhm.ac.uk/research-curation/scientific-resources/biodiversity/uk-biodiversity/uk-species/about-the-species-inventory/index.html>.
- United States Environmental Protection Agency. 2014.** Tutorials on Systems Thinking using the DPSIR Framework. Available at: <http://www.epa.gov/ged/tutorial/>.
- US Environmental Protection Agency. 1975.** *DDT: A ! review of scientific and economic aspects of the decision to ban its use as a pesticide*. EPA-540/1-75-022: Washington DC.
- Van Swaay CAM, Van Strien AJ, Harpke A, Fontaine B, Stefanescu C, Roy D, Maes D, Kühn E, Ōunap E, Regan E, Švitra G, Heliölä J, Settele J, Warren MS, Plattner M, Kuussaari M, Cornish N, Garcia Pereira P, Leopold P, Feldmann R, Jullard R, Verovnik R, Popov S, Brereton T, Gmelig Meyling A, Collins S. 2010.** The European Butterfly Indicator for Grassland species 1990–2009. Report VS2010.010, De Vlinderstichting, Wageningen.
- VanDerWal J, Shoo LP, Johnson CN, Williams SE. 2009.** Abundance and the environmental niche: environmental suitability estimated from niche models predicts the upper limit of local abundance. *The American Naturalist* **174**: 282–291.
- Wallis De Vries MF, Baxter W, Van Vliet AJH. 2011.** Beyond climate envelopes: effects of weather on regional population trends in butterflies. *Oecologia* **167**: 559–571.
- White P, Kerr J. 2006.** Contrasting spatial and temporal global change impacts on butterfly species richness during the 20th century. *Ecography* **29**: 908–918.
- Wiens JA, Rotenberry JT, Van Horne B. 1986.** A lesson in the limitations of field experiments: shrubsteppe birds and habitat alteration. *Ecology* **67**: 365–376.