LAND AND WATER
www.csiro.au

Checking for Change

The science behind practical monitoring of ecological improvement

Erik Doerr, Veronica Doerr, Jacqui Stol, and Micah Davies

Citation

Doerr E, Doerr V, Stol J, Davies M and (2017) Checking for Change: the science behind practical monitoring of ecological improvement. CSIRO, Australia.

This research report, the practical monitoring guide produced as a result of this research, and associated indicator factsheets, datasheets and instructional videos are all available at the 'Checking for Change' BioCollect website, where data can also be entered and stored. Visit: http://tinyurl.com/checking4change

Acknowledgements

We sincerely appreciate the generous contributions of all the land owners and land managers on whose properties this research was undertaken. Many landowners also provided valuable feedback in the development of the practical guide and we are very grateful to them all. Multiple natural resource managers and staff across a wide range of organisations provided insightful comments, expert observations, and peer review of this work, including the Department of the Environment and Energy, Greening Australia, NSW Local Land Services, NSW Office of Environment, the Nature Conservation Trust and Landcare. The original research, as well as the development of the resulting practical guide, were assisted by the New South Wales Government through its Environmental Trust. We are also grateful to Sue McIntyre, Saul Cunningham for scientific peer review of this work.

Contents

Introduction

This report is intended as a companion to the monitoring guide 'Checking for Change: a practical guide to checking whether sites newly managed for conservation are on track to improve' (Stol et al. 2016), available at http://tinyurl.com/checking4change, and supported by the Atlas of Living Australia citizen science portal BioCollect (www.ala.org.au/biocollect). At the time of publication of this report, manuscripts for scientific journals were in preparation and review and not yet publicly available. Thus, this report has been developed to give users of the monitoring guide direct access to a description of the original research used to develop this new monitoring approach.

Ecological condition can be defined as the degree to which an ecosystem exhibits its full complement of composition, structure and function at genetic, species and ecosystem scales (Noss 1990, Gibbons and Freudenberger 2006, Stoddard et al. 2006). Increasingly, environmental managers are trying to improve the condition of previously degraded areas based on this definition. However, our current state of knowledge about which management actions are most likely to achieve this is incomplete and insufficient to predict the exact consequences of our interventions (see https://research.csiro.au/biodiversity-knowledge/projects/knowledge-bank/). Adaptive management (Holling 1978) is frequently advocated as a solution to encourage 'learning by doing' and thus modify management approaches as necessary to eventually achieve successful outcomes across these levels and scales. However, successful adaptive management requires the use of effective monitoring programs that can detect changes in critical ecological parameters (Noss 1990, Lindenmayer et al. 2012).

A recent review by Westgate et al. (2013) found that adaptive management projects were often of short duration and were hindered by the expense and difficulties of long-term monitoring. Thus, while much has been written about what makes a good ecological indicator (e.g., Noss 1990, Gibbons and Freudenberger 2006), one important criterion that has been somewhat overlooked is the ability of an indicator to register an informative change over short time intervals. The more rapidly we can detect change in the condition of a site, the more we can implement adaptive management (Lindenmayer et al. 2012). Furthermore, as conservation efforts increasingly shift towards private lands (Stolton et al. 2014), there is a need for indicators that are measurable by land owners and not just ecological experts (Schulze et al. 2009).

To help address these gaps in useful ecological monitoring indicators, we developed a suite of 24 candidate indicators (plus one composite metric) and tested whether they responded to changes in management over short time frames in temperate eucalypt woodland sites in south-eastern Australia. Our candidate indicators were derived from existing indicators as well as novel indicators that we developed. For both existing and novel indicators, we particularly focused on surrogates of ecosystem function and on indicators that could potentially be assessed by most land managers. We monitored these indicators at 20 stewardship sites (where livestock had been removed within the previous few years) and at 20 matched control sites (where livestock grazing continued) in the springs of 2009 and 2011, and then re-assessed most of the indicators again in 2015. Subsequent analyses allowed us to identify a suite of indicators that were capable of detecting the effects of recent management change (i.e., livestock removal) over a 2-6 year time interval. We also assessed the accuracy and ease with which landholders themselves could use the indicators to collect meaningful data.

Based on the results of these analyses, we generally recommended that monitoring programs should focus more on the ground layer (including down into the soil) and organisms with short generation times. We also suggested that measures of ecosystem function in particular may be useful for developing indicators specifically for adaptive management across a broad suite of ecological communities. We used the results to develop our practical monitoring guide 'Checking for Change' to allow landholders to make use of this research to monitor their own properties as well as to allow more widespread yet consistent data collection to help with regional, state and national reporting and adaptive management of environmental improvement programs.

Methods

Developing potential indicators

To identify potential indicators for adaptive management, we articulated the criteria they would need to meet in order to be successfully used. We reviewed a wide range of existing indicators and selected a subset for testing that met our criteria. We also modified existing indicators to make them more suitable for use by non-experts. Finally, we developed a number of new, novel indicators that met our criteria in order to have a broad suite to test in this study (Doerr et al. 2012). Criteria for inclusion in the list of potential indicators were:

Sensitive over short time frames

To decide which indicators to test, we only considered those that logically had the potential to be sensitive (show change) within <5 years. For example, # of mature trees does not have the potential to change in a measurable, statistically analysable way over short time frames so was not included as a potential indicator.

Improvement, not just change

To be included as a potential indicator, it was important to have enough prior information to establish that change in the parameter being measured truly indicates improvement, not just change. For each potential indicator, we needed to have confidence that we could specify the direction of change that indicates better or worse ecological condition (and/or what level might constitute an optimum). In most cases, we selected potential indicators that were previously shown through published research to differ between reference or benchmark sites and degraded sites. Research then focused on testing whether management actions on degraded sites yielded a measurable return toward reference conditions over short time frames.

We also needed to avoid testing indicators that are potentially *only* short-lived, with no longer-term consequences for recovery of a site. Therefore data were analysed at both two and six years after a formal change to conservation management. However, some short-term responses may be important precursors for later changes and are thus meaningful signs of improvement even if they do not continue to improve beyond the first years after a change in management. We developed a simple theoretical model of the process of ecological recovery, involving three main ways in which different aspects of ecological condition may respond over time and eventually lead to full and long-term improvement in overall ecological condition (Figure 1).

In this model, Type I indicators measure aspects of the ecosystem that respond quickly, and their improvement creates the necessary conditions for Type II indicators to begin showing a response. Thus, note that in the figure, Type II indicators begin to show significant improvement only after Type I indicators fully improve. In selecting potential indicators to test, we paid particular attention to whether they might be such Type I indicators – measuring aspects of the environment that need to improve first in order to stimulate further recovery. For example, changes in the soil layer may be important precursors that facilitate change in vegetation, so were prioritised as potential indicators to test.

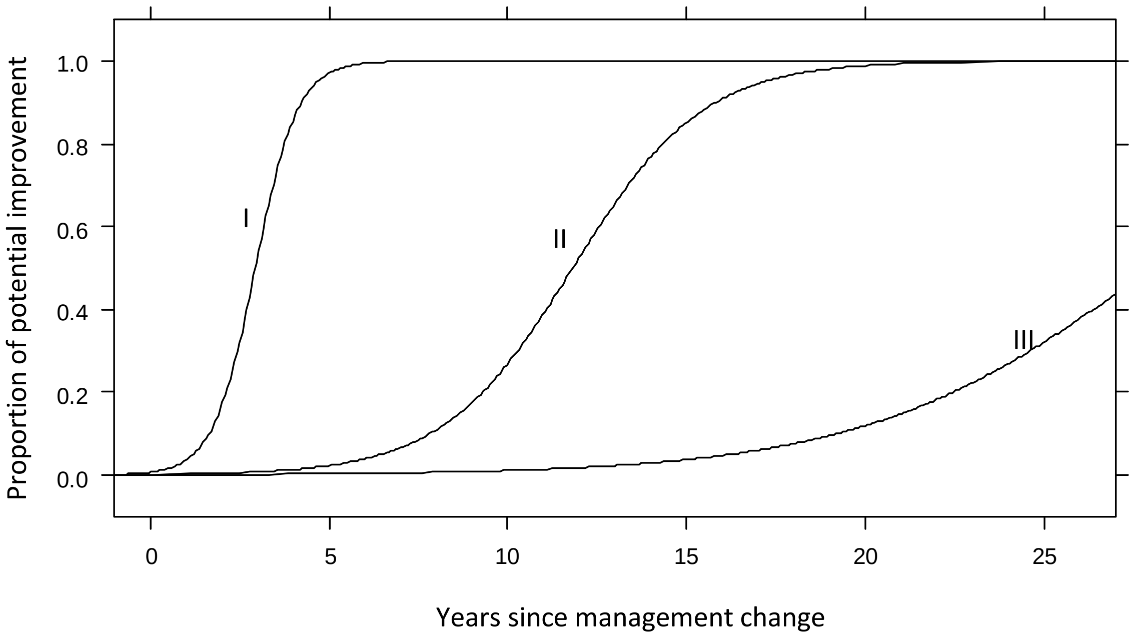

Figure 1. Three hypothesised types of indicators based on their functional response trajectories over time. Note that Type II indicators in particular may be dependent on recovery of Type I indicators to create the necessary preconditions for their own improvement.

Rapid, simple and inexpensive to measure

We only evaluated indicators that had the potential to meet these criteria - that were relatively quick, easy and inexpensive to measure. Thus, indicators involving the use of deep specialist knowledge or equipment were not included (e.g. those that required plant species identification or $1m^2$ quadrats were not assessed). We also explored whether simpler, non-specialist alternatives could be developed by modifying existing indicators and then testing them. For example, we modified some traditional indicators to use 'morphospecies' approaches to determine apparent species richness rather than relying on detailed species identification, particularly where existing scientific literature suggested that morphospecies richness might be a useful surrogate.

Direct relationship to composition, structure and especially function

We also aimed to test potential indicators across all three aspects of condition: composition, structure and function. In general, there are few indicators currently available for function, yet our research hypothesis was that that functions are more likely to be Type I indicators – aspects of the system that need to improve in order for other traits (like those related to composition and structure) to improve. Thus, we developed our own indicators of function (or structural surrogates of function). These included simple measures of attributes such as pollination rates and litter decomposition rates. Although these were novel indicators, they were very much founded on evidence-based scientific research showing that they are likely to be measures of improvement, not just change.

The final set of potential indicators we tested in this study is shown in Table 1. It includes one composite measure, the site score from the Conservation Value Measure developed to support the Australian government's investment choices in incentive-based stewardship programs for Box Gum Grassy Woodlands. This is because this measure was actively being used as a measure of condition, such composite metrics were increasingly being advocated by government as monitoring indicators, and we wanted to be able to evaluate its performance relative to individual indicators. We specifically hypothesised that individual indicators might meet the criteria for supporting adaptive management more closely than a composite metric could, in which case it could be counter-productive to advocate for composite metrics at the expense of individual indicators.

Table 1. List of potential indicators tested, with short names and information on indicator type (C = Composition; S = Structure; F = Function), whether the variable is already commonly used as an indicator, whether measurement of the variable requires some expert knowledge, how to interpret a change in the indicator in terms of ecological condition, and references which provide either a theoretical basis for the indicator or an example of its use in monitoring.

Variable name	Short name	Type	Common?	Expert?	Interpretation	References
Plant morphospecies	Pmorph	C	No	No	more = better	Schulze et al. 2009
Native plant morphospecies	NPmorph	C	No	No	more = better	Schulze et al. 2009
Lepidoptera morphospecies	LepMorph	C	No	No	more = better	Lomov et al. 2006; Rakosy and Schmitt 2011
Lepidoptera diversity (Shannon index)	ShanLep	C	No	No	higher = better	Lomov et al. 2006; Rakosy and Schmitt 2011
Ant mound entrances	AntEnts	C	No	No	more = better	Andersen & Majer 2004; Underwood & Fisher 2006
Litter invertebrates	LitInverts	C	No	No	more = better	Lindsay & Cunningham 2009
Bird species richness	BirdSpp	C	Yes	Yes	more = better	Henry et al. 2008; Eglington et al. 2012
Bird diversity (Shannon index)	ShanBird	C	Yes	Yes	higher = better	Henry et al. 2008; Eglington et al. 2012
Optimal native perennial basal cover	OptNPBC	S	Yes	Yes	optimum = 15%	Spooner & Briggs 2008; Gibbons et al. 2009
Optimal native perennial foliage cover	OptNPFC	S	Yes	Yes	optimum = 70%	Spooner & Briggs 2008; Gibbons et al. 2009
Traditional litter cover	TradLitCov	S	Yes	No	higher = better	Parkes et al. 2003; Price et al. 2010
Alternative litter cover	AltLitCov	S(F)	No	No	higher = better	Parkes et al. 2003; Price et al. 2010
Litter depth	LitDepth	S(F)	Yes	No	deeper = better	Bugalho et al. 2011
Bare ground cover	BGCov	S(F)	Yes	No	lower = better	Schulze et al. 2009; Price et al. 2010
Optimal mean interperennial distance	OptMn-IPD	S(F)	No	Yes	optimum = 150mm	Lindsay & Cunningham 2012
Optimal SD interperennial distance	OptSD-IPD	S(F)	No	Yes	optimum = 130mm	E.A. Lindsay pers. comm.
Regeneration (yes/no)	Regen1	F	Yes	No	yes = better	Gibbons et al. 2009
Regeneration (# seedlings <1m)	Regen2	F	Yes	No	more = better	Spooner & Briggs 2008
Flowers	Flowers	F	No	No	more = better	Oliver 2002
Seed stems	Seeds	F	No	No	more = better	Oliver 2002
Pollinators	Polls	F	No	No	more = better	Lomov et al. 2010
Litter (silverbeet) decomposition rate	SBDecomp	F	No	No	higher = better	Lindsay & Cunningham 2009
Litter decomposition score	LitDecomp	F	No	No	higher = better	Tongway and Hindley 2004
Optimum soil coherence	OptCoher	F	No	No	higher = better	Tongway and Hindley 2004
Conservation Value Measure Site Score	CVM SS	All	Yes	Yes	Optimum = reference value	Gibbons & Ryan 2008

While most of these indicators are largely self-explanatory and more detail can be found in Doerr et al. (2012), a few warrant a short explanation. 'Litter invertebrates' refers to a quick count of the abundance of invertebrates, regardless of taxa, observed on the surface, assessed through simply disturbing the litter and counting invertebrates seen. 'Alternative litter cover' suggests that the absolute amount of litter cover is not as important as whether areas not occupied by plants contain litter (as opposed to bare ground, rock, etc.), and thus is a measure of the proportion of ground *without a plant* that has litter. 'Optimal mean interperennial distance' and 'Optimal Standard Deviation (SD) interperennial distance' aim to explore the need for inter-tussock spaces in Australian grassy ground layers (to provide spaces for forbs and animal foraging) and some heterogeneity or patchiness in those spaces. These potential indicators involve measuring distances between the edges of plant butts and comparing them to optimal values derived from reference grassy woodland sites. Finally, 'Litter (silverbeet) decomposition rate' aimed to explore whether it was possible to assess litter decomposition rates directly but rapidly, by measuring the change in surface area of dried silverbeet leaves placed in the leaf litter for a few months. Silverbeet leaves were used because native eucalypt leaves naturally decompose at slow rates making them impractical for rapid assessment.

Experimental design for assessing improvement

We identified 20 stewardship sites where a change from grazing to conservation management had been recently implemented in grassy woodlands of south-eastern Australia. These were paired with 20 control sites (typically on an adjoining property or adjacent paddocks) where the ecosystem type, dominant species and overall floristic structure were similar but livestock grazing for production purposes was intended to continue. Properties were located in New South Wales between Blayney in the north, and Murrumbateman near Yass in the south.

Baseline data on our 25 potential indicators (Table 1) were collected in spring 2009 at both stewardship sites and control sites. Sampling was repeated 2 years later (spring 2011) at exactly the same locations to facilitate robust statistical data to show if ecological indicators demonstrated any significant improvement in stewardship sites compared with control sites, while accounting for a range of possible confounding factors (such as differences in rainfall between years or across the study region). We repeated data collection for 17 of the indicators in spring 2015, 6 years after the first surveys. The indicators that were not included in 2015 were those that had showed no promise due to logistical problems in the 2009-2011 analyses (e.g. litter decomposition rate was not re-evaluated because of disturbance by livestock in 2009 and 2011 resulting in a lack of analysable data).

Data were collected along two 50m transects and 12 1m^2 quadrats that were randomly placed within each of the 40 sites. Data were collected at four scales – whole-of-site, whole-of-transect, quadrat, and point depending on the nature of the variable being quantified. For example, Lepidoptera (butterfly and moth) morphospecies was assessed using whole-of-site surveys, Regeneration (# seedlings<1m) was counted at the whole-of-transect scale, the number of Flowers was counted at the quadrat scale, and groundcover variables were assessed at the point scale using the four corners plus the centre of each quadrat. Data were then combined using averages or density estimates as appropriate to create whole-of-site estimates for all variables.

Analyses then focused on whether changes over time at stewardship sites represented improvement *relative to* changes over time at control sites. This allowed us to evaluate improvement without specifically having to measure the many other factors that could influence what is observed at a stewardship site. For example, ecological condition could be declining over a whole region due to widespread threats or the delayed effects of past land clearing, in which case stewardship sites might decline in condition but do so at a slower rate than control sites. This form of analysis was designed to detect those situations, as well as overall improvement in stewardship sites, because both provide evidence that site-scale management is effective, at least in terms of addressing site-scale management challenges. More information on the types of statistical analyses performed at the different time periods is provided in the following sections.

Analysing improvement after two years

Note that while many of our indicators were naturally correlated, this did not present a problem for our analyses as we evaluated the performance of each potential indicator in separate analyses rather than analyse the optimum set of indicators. For example, the two regeneration indicators are certainly correlated, but we were specifically interested in which one (if either) would be better able to show short-term improvement.

We used the 2011 data for each indicator as a response variable and the 2009 data as one of our predictors (the 'Baseline'), and employed a model-selection process based on that described by Zuur et al. (2009). The distribution of the data for many variables violated assumptions of statistical analyses, so we used a range of data transformations to enable the data to be suitable for analysis, including log, square root and logit transformations. We performed all subsets regression for each indicator using four site covariates as candidate predictor variables (mean annual rainfall, mean annual temperature, woodland structure, and intensity of historical livestock grazing). We used the 2009 baseline data as a forced predictor (i.e., always included) and selected as the best model the one with the lowest Akaike Information Criteria scores corrected for small sample size (AICC).

We then performed analysis of covariance (ANCOVA) to test for the main treatment effect of interest (that of the binary Stewardship variable) using the variables from this best model as covariates. We repeated ANCOVA without site covariates (including only the 2009 baseline as a covariate) and finally performed a regression analysis including only the 2009 baseline as a predictor. This process resulted in four variously nested models for comparison: the "Full model" containing the 2009 baseline, the Stewardship variable, and covariates (those from the best covariate model); the "Baseline + covariates model" (the best covariate model); the "Baseline + Stewardship model"; and the "Baseline only model". We calculated AICC scores and Akaike weights for each of these four models. Significance of treatment effects (i.e. whether stewardship sites improved relative to control sites) were assessed based on the Akaike weights of models containing the stewardship variable, and on the p-values in these models. All analyses were carried out using SYSTAT 13 (SYSTAT Software, Inc.).

Analysing improvement after six years

Analysing improvement after six years required a somewhat different approach for a few reasons. First, the additional time step (data collection in 2009, 2011, and 2015) now created a time series of data rather than just two points in time for each site. Second, the distribution of the data collected changed such that in many cases, data transformations were no longer able to ensure that assumptions of the statistical methods used above would not be violated. As a result, improvement after six years was analysed using non-parametric Wilcoxon sum rank tests to test for differences between stewardship and control sites in the *change* in indicator values over various time periods. In other words, the difference in indicator values between 2009 and 2015 was calculated for each site, then these values for stewardship versus control sites were compared. The same approach was used to re-analyse the comparisons between 2009 and 2011 for consistency. All analyses were one-tailed as our hypotheses involved clear predictions about the direction of difference we expected.

One disadvantage of this non-parametric approach was that we were unable to include covariates in these analyses, so we were unable to account for variation in stewardship improvement that might be due to regional conditions or prior land management. Indeed, covariates were often part of the best final models in the analyses of improvement after two years. Thus, there is likely to be greater noise in the analyses of improvement after six years, making it harder to detect any positive effects of management that exist in reality. As a result, decisions about the final set of most useful indicators were made using a combination of results from the initial two-year analyses, results from the non-parametric comparisons after two and six years, and visual inspection of box plots created to show median, quartile and outlier values for each indicator for stewardship versus control sites in 2009, 2011, and 2015.

Evaluating accuracy and likelihood of adoption by land managers

In addition to analysing improvement in the indicators as calculated by our research team, we also worked with land managers to less formally evaluate which indicators were likely to be useful to them and actually adopted. We designed land manager surveys to provide insight into: 1) the ability of land managers to accurately quantify each candidate indicator, 2) the degree to which land managers found the methods simple and easy, and 3) the degree to which land managers believed the indicators were likely to provide reliable information on short-term change.

To limit the time required of land managers who volunteered to help with this process, we only designed the survey to evaluate 16 of the 25 candidate indicators – those that were most feasible to assess with land managers in a single survey without additional mathematical calculations. As these land manager surveys were conducted concurrently with ecological data collection in 2011, choice of indicators to assess was not yet informed by any ecological data analyses. Thus, some indicators that subsequently proved infeasible for ecological reasons were still assessed.

The survey was designed in two parts – Part 1 consisted of a series of questions to assess the simplicity and usefulness of each indicator (and corresponding method) as perceived by the land manager. These were designed to be neutral, non-leading questions to elicit answers that were categorical and could be allocated a percentage response value (e.g. 75% of respondents found Survey Method 1 to be 'Very Simple'). The second part of the survey was designed as five broader, open ended questions which are appropriate for surveying attitudes, feelings and opinions (Iarossi 2006). These resulted in less quantifiable responses but provided valuable feedback on many issues including *why* land managers found different methods to be useful, whether they believed they had gained new skills, and how the methods could be improved. Full details on the survey can be found in Doerr et al. (2012).

Fourteen land managers from ten of the project's 20 stewardship sites trialled the indicators and then completed the survey. All land managers were given the same background information, including a handout providing a description of the methods for measuring each indicator and a brief scientific justification of why we suspected it might be useful. Minimal additional verbal instructions were provided to mimic the experience of using the indicators based on a written guide. The ecological data recorded by land managers during this process were also compared with data collected simultaneously by one member of our research team (JS) to assess accuracy from the researcher perspective not just the landholder perspective. Results were qualitatively analysed and used to inform the final set of indicators considered to be most useful for supporting adaptive management.

Results

Improvement after two years

A number of the indicators revealed statistically significant improvement at stewardship sites compared to control sites, even when controlling for covariates like recent annual rainfall in parametric analyses (Table 2). Our indicators for litter decomposition, abundance of litter-dwelling invertebrates, litter depth, and percent cover of bare ground exhibited clear improvement over a two-year interval at stewardship sites compared to control sites. Five other indicators – interperennial distances, litter cover, foliage cover of native perennial plants, number of native plant morphospecies, and the composite metric (the site score from the Conservation Value Measure) – showed trends of positive change (particularly for interperennial distances and litter cover) which may become significant over a somewhat longer time period. The other eleven indicators analysed showed no signs of change over the two years of the study. Although it is possible above average annual rainfalls during the survey period maximised our chance of observing change, we specifically undertook statistical analysis to detect changes at these sites relative to control sites. In other words, the positive effect of stewardship management could not simply be an artefact of above average rainfalls, though the speed with which these changes occur enough to be detectable may vary depending on recent local conditions.

Similar, but not identical, patterns were observed when using non-parametric analyses of the same structure used to analyse the improvement after six years (Table 2).

Improvement after six years

Not surprisingly, a greater number of indicators revealed statistically significant improvement at stewardship sites compared to control sites after six years had passed since the change in management instead of just two years, despite the use of simpler, non-parametric analyses (Table 2). Plant morphospecies improved (both native and in general, without requiring specialist knowledge of native species), as did bird species richness, both measures of native perennial cover, litter depth, bare ground (reduced), the level of litter decomposition, and the composite CVM site score.

Perhaps surprisingly, a few of the indicators that showed promise after two years were no longer showing improvement at six years. Litter invertebrates in particular had looked very promising at two years but that did not translate into clear, lasting improvement at six years after the change in management. This may partly be because it targets something that can change very rapidly both up and down, and 'improvement' may actually involve very episodic pulses of invertebrate activity which then quickly translate into other improvements in the soil and vegetation (assessed by Type II indicators). It is worth continuing to explore this possibility.

Ease and accuracy of use by land managers

Overall, the majority of land managers found most of the survey methods either 'very simple' or 'simple' as well as useful for detecting change in the ecological condition of a site. Some of the indicators we initially assumed would be more difficult for land managers to use were actually ones they believed were simple and they demonstrated confidence in using, like counting plant morphospecies. This was despite the fact that most land managers could not or were not confident with identifying actual plant species. Similarly, some of the indicators we assumed would be easy to use and that land managers would immediately perceive as useful ended up ranking low in terms of overall land manager friendliness, such as the proportion of bare ground. Interestingly, many of our novel indicators based on function rather than composition or structure were actually deemed simple and useful by land managers despite the fact that they are relatively new and not in common practice at the moment. The accuracy of these data collected by land managers was comparable to that collected by an ecological expert for most of the indicators. As a result of the land manager testing, we found many planned to use a selection of these newly learned techniques to assist future monitoring and understanding of trends in ecological change. Full results of the land manager assessments are presented in Doerr et al. (2012) and a summary of overall 'land manager friendliness' based on simplicity, usefulness, and accuracy of data collection is presented in Table 2.

Final set of most useful indicators

We used a combination of all these results to select a final set of indicators to advocate for short-term assessment of improvement in ecological condition to support adaptive management. The choice of final indicators was made by qualitative assessment of the full set of results (see final two columns in Table 2). Indicators that were both statistically significant indicators of change over both two and six years and were also rated as land manager friendly were obvious choices to include. Some indicators were chosen because the statistical analyses suggested they were useful indicators but they weren't as land manager friendly as others. In these cases, we believed that better quality instructions and simpler methodology could be developed to make them more land manager friendly.

No indicators were selected purely because they were land manager friendly – all needed to show statistical evidence of reliably revealing short-term improvements in ecological condition. The only exception was Litter invertebrates. As noted above, this indicator showed statistically significant improvement in stewardship sites compared to control sites after just two years using our parametric analyses with covariates, but the improvement was no longer evident at six years (nor in the non-parametric analyses for the two-year comparison). Annual rainfall was the most important covariate in the final parametric model for this indicator. These results suggest that there may be something interesting happening under higher rainfall conditions in terms of stimulating short-duration processes important for longer term recovery. This possibility of episodic 'pulses' of recovery accords well with our current general understanding of the dynamics of Australian temperate and sub-tropical ecosystems. As this indicator was also one of the most land manager friendly, it may be worth advocating as an indicator that we need to learn more about, and can with the assistance of land managers themselves.

Table 2. List of potential indicators with summary of results from two-year analyses (2009-2011), six-year analyses (2009-2015), and land manager surveys, with final decision and notes on usefulness as part of a set of indicators to support adaptive management. 'Yes' indicates significant difference in improvement between stewardship and control site with $p<0.05$, 'probably' indicates difference in improvement with $p<0.10$, 'maybe' indicates best model includes the variable but $p>0.10$, 'no' indicates no statistical evidence for difference in improvement between stewardship and control sites, and a dash indicates no analysis performed. Scores for 'Land manager friendly?' range from 0 to 3 and are based on receiving one point each if rated as simple by land managers, one point if rated as likely to show improvement by land managers, and one point if accurate data collected by land managers compared to member of research team. Thus, the higher the score, the more land manager friendly the indicator is.

Variable name	Rejected-infeasible	2-yr improvement (parametric)	2-yr improvement (non-parametric)	6-yr improvement (non-parametric)	Land manager friendly?	Final set?	Selected Notes
Plant morphospecies		no	no	yes	3	yes	
Native plant morphospecies		maybe	yes	yes	--	yes	A stronger indicator than simply plant morphospecies, but one that requires some specialist knowledge. Land managers could choose which morphospecies indicator to use based on self-assessment of their ability to recognise native plants
Lepidoptera morphospecies		no	no	no	2	no	
Lepidoptera diversity		no	no	no	--	no	
Ant mound entrances	rejected	--	--	--	2	no	Infeasible because too few were detected, resulting in a majority of 0 values in the data
Litter invertebrates		yes	no	no	3	yes	Only short-term evidence of improvement but managers learned the most from it – recommended as one to experiment with rather than definitely indicate improvement
Bird species richness		no	no	yes	--	yes	Requires some specialist knowledge but worth using a morphospecies approach
Bird diversity		no	no	no	--	no	

Variable name	Rejected-infeasible	2-yr improvement (parametric)	2-yr improvement (non-parametric)	6-yr improvement (non-parametric)	Land manager friendly?	Final set?	Selected Notes
Optimal native perennial basal cover		no	no	yes	3	yes	
Optimal native perennial foliage cover		maybe	yes	yes	1	yes	Less land manager friendly than basal cover, but shows a quicker response so worth suggesting both
Traditional litter cover		no	no	no	--	no	
Alternative litter cover		probably	no	maybe	--	no	
Litter depth		yes	maybe	yes	1	yes	While not particularly land manager friendly, this was a reliable early indicator so better instructions/methods were provided to make it more land manager friendly
Bare ground cover		yes	no	yes	1	yes	While not particularly land manager friendly, this was a reliable early indicator so better instructions/methods were provided to make it more land manager friendly
Optimal mean interperennial distance	partially rejected	probably	--	--	3	no	Not assessed at six years because required assessment in autumn rather than spring, and thus not land manager friendly as part of a monitoring package, given that other indicators need to be assessed in spring
Optimal SD interperennial distance	partially rejected	no	--	--	--	no	Not assessed at six years because required assessment in autumn rather than spring, and thus not land manager friendly as part of a monitoring package, given that other indicators need to be assessed in spring
Regeneration (yes/no)		no	no	no	--	no	

Variable name	Rejected-infeasible	2-yr improvement (parametric)	2-yr improvement (non-parametric)	6-yr improvement (non-parametric)	Land manager friendly?	Final set?	Selected Notes
Regeneration (# seedlings <1m)	partially rejected	no	--	--	3	no	Not assessed at six years because most sites had 0 so the binary (yes/no) analysis made better use of the data
Flowers	rejected	--	--	--	2	no	Infeasible as almost all flowers were from non-native species
Seed stems	rejected	--	--	--	0	no	Infeasible as almost all seed stems were from non-native species
Pollinators	rejected	--	--	--	2	no	Infeasible as pollinators were very rarely detected, resulting in a majority of 0 values in the data
Litter (silverbeet) decomposition rate	rejected	--	--	--	2	no	Infeasible as livestock regularly disturbed the silverbeet despite specific attempts to avoid such disturbance. It is worth noting that 'tea bag decomposition' (Keuskamp et al. 2013) is currently under development as a promising alternative
Litter decomposition score		yes	yes	yes	2	yes	
Optimum soil coherence	partially rejected	no	--	--	2	no	Not assessed at six years because even the ecological experts did not feel confident in using the scoring system
Conservation Value Measure Site Score		maybe	no	yes	--	no	While a reliable indicator in the medium term, such a metric is impractical for rapid assessment by land managers themselves – included in our research as a point of comparison

Discussion

Learning about ecological recovery – processes and functions

In general these results lend support to our simple theory of ecological recovery (Figure 1) as some indicators showed early response to improved management, and were logically or theoretically ‘prerequisites’ for improvement in other factors. Not surprisingly, most of these early, ‘Type I’ indicators were aspects of the soil and ground layer, which are perhaps both more able to respond rapidly and of foundational importance to resident plants and animals and shape a full biodiverse, high functioning site.

However, conducting this research also helped reinforce that researchers actually know little about the *process* of ecosystem recovery - i.e. *how* it happens in practice. We often count and catalogue what’s at a site and compare across sites rather than deeply research processes of reproduction, establishment, soil structure formation, etc. Similarly, we frequently study what happens when a site degrades due to various threats, but removing threats does not guarantee that processes are in place to allow recovery. This lack of knowledge about the process of recovery is partly because on-ground attempts to recover native ecosystems are only relatively recent, and full recovery (assuming it is possible) will likely take decades.

It is also worth noting that with disruptors such as global climate change, ideal ‘reference’ conditions are expected to change as new bioclimatic conditions at any given site will make it suitable for a different suite of species and thus a different ecosystem (Williams et al. 2014). These shifting dynamics of ecosystems have always occurred but on relatively long time scales, largely irrelevant to our management. The velocity of climate change will increasingly pressure ecosystems to change more rapidly, making these dynamic changes much more relevant to management. As a result, measures of composition and structure may become much less reliable indicators of ‘improvement’. We will increasingly need to develop reliable function and process indicators that reveal something about the ability of the ecosystem to respond to changing environmental conditions, or aspects of ecosystem health that transcend specific composition and structure (Prober et al. 2015). Some of the indicators we have explored in this research may be useful in this new context (e.g. litter decomposition or optimal cover) but others (e.g. bird species richness) will potentially be replaced by better indicators of ecosystem and landscape function as new empirical research is directed toward better understanding and managing ‘dynamic ecosystems’ and ‘dynamic nature’.

From research to a practical monitoring guide

Many of our final indicators are not currently in use by monitoring programs in Australia, particularly not in the often simplified, land manager friendly forms in which we explored them. Additionally a number of indicators that *are* currently in use failed to show improvements within 2-6 years in this study. This suggests that monitoring of ecological condition specifically to inform adaptive management could be improved, and approaches to do this should focus on our final short-term indicators which could usefully complement longer-term monitoring to understand the full dynamics of recovery over decades. Thus, we developed the guide ‘Checking for Change: a practical guide to checking whether sites newly managed for conservation are on track to improve‘ (Stol et al. 2016), as well as a range of supporting resources (all available at http://tinyurl.com/checking4change) to ensure our final indicators were widely accessible.

References

Andersen, A. N., and J. D. Majer. 2004. Ants show the way Down Under: invertebrates as bioindicators in land management. Frontiers in Ecology and the Environment **2**:291-298.

Bugalho, M. N., X. Lecomte, M. Goncalves, M. C. Caldeira, and M. Branco. 2011. Establishing grazing and grazing-excluded patches increases plant and invertebrate diversity in a Mediterranean oak woodland. Forest Ecology and Management **261**:2133-2139.

Doerr, E. D., J. Stol, M. J. Davies, and V. A. J. Doerr. 2012. Linking Incentives to Outcomes for NRM. Final report to the New South Wales Environmental Trust. CSIRO Ecosystem Sciences, Canberra, Australia. *Available from the Environmental Trust upon request.*

Eglington, S. M., D. G. Noble, and R. J. Fuller. 2012. A meta-analysis of spatial relationships in species richness across taxa: Birds as indicators of wider biodiversity in temperate regions. Journal for Nature Conservation **20**:301-309.

Gibbons, P., S. V. Briggs, D. Ayers, J. Seddon, S. Doyle, P. Cosier, C. McElhinny, V. Pelly, and K. Roberts. 2009. An operational method to assess impacts of land clearing on terrestrial biodiversity. Ecological Indicators **9**:26-40.

Gibbons, P., and D. Freudenberger. 2006. An overview of methods used to assess vegetation condition at the scale of the site. Ecological Management & Restoration **7**:S10-S17.

Gibbons, P., and P. Ryan. 2008. A conservation value index for Box Gum Grassy Woodland. Final Report to Accompany Version 1.6. The Fenner School of Environment and Society, The Australian National University, Canberra.

Henry, P. Y., S. Lengyel, P. Nowicki, R. Julliard, J. Clobert, T. Čelik, B. Gruber, D. S. Schmeller, V. Babij, and K. Henle. 2008. Integrating ongoing biodiversity monitoring: Potential benefits and methods. Biodiversity and Conservation **17**:3357-3382.

Holling, C. S. 1978. Adaptive environmental assessment and management. John Wiley and Sons, New York.

Iarossi, G. 2006. The power of survey design: a users guide for managing surveys, interpreting results, and influencing respondents. The World Bank, Washington D.C.

Keuskamp J. A., Bas, J., Dingemans, J., Lehtinen, T., Sarneel, J. M. and Hefting, M. M. 2013. Tea Bag Index: a novel approach to collect uniform decomposition data across ecosystems Methods in Ecology and Evolution **4**:1070–1075.

Lindenmayer, D. B., P. Gibbons, M. Bourke, M. Burgman, C. R. Dickman, S. Ferrier, J. Fitzsimons, D. Freudenberger, S. T. Garnett, C. Groves, R. J. Hobbs, R. T. Kingsford, C. Krebs, S. Legge, A. J. Lowe, R. McLean, J. Montambault, H. P. Possingham, J. Radford, D. Robinson, L. Smallbone, D. Thomas, T. Varcoe, M. Vardon, G. Wardle, J. Woinarski, and A. Zerger. 2012. Improving biodiversity monitoring. Austral Ecology **37**:285-294.

Lindsay, E. A., and S. A. Cunningham. 2009. Livestock grazing exclusion and microhabitat variation affect invertebrates and litter decomposition rates in woodland remnants. Forest Ecology and Management **258**:178-187.

Lindsay, E. A., and S. A. Cunningham. 2012. Effects of exotic grass invasion on spatial heterogeneity in the ground-layer of grassy woodlands. Biological Invasions **14**:203-213.

Lomov, B., D. A. Keith, D. R. Britton, and D. F. Hochuli. 2006. Are butterflies and moths useful indicators for restoration monitoring? A pilot study in Sydney's Cumberland Plain Woodland. Ecological Management & Restoration **7**:204-210.

Lomov, B., D. A. Keith, and D. F. Hochuli. 2010. Pollination and plant reproductive success in restored urban landscapes dominated by a pervasive exotic pollinator. Landscape and Urban Planning **96**:232-239.

Noss, R. F. 1990. Indicators for Monitoring Biodiversity - a Hierarchical Approach. Conservation Biology **4**:355-364.

Oliver, I. 2002. An expert panel-based approach to the assessment of vegetation condition within the context of biodiversity conservation: Stage 1: the identification of condition indicators. Ecological Indicators **2**:223-237.

Parkes, D., G. Newell, and D. Cheal. 2003. Assessing the quality of native vegetation: The 'habitat hectares' approach. Ecological Management & Restoration **4**:S29-S38.

Price, J. N., N. K. Wong, and J. W. Morgan. 2010. Recovery of understorey vegetation after release from a long history of sheep grazing in a herb-rich woodland. Austral Ecology **35**:505-514.

Prober, S. M., K. J. Williams, T. D. Harwood, V. A. J. Doerr, T. Jeanneret, G. Manion, and S. Ferrier. 2015. Helping Biodiversity Adapt: Supporting climate adaptation planning using a community-level modelling approach. CSIRO Land and Water, Canberra.

Rákosy, L., and T. Schmitt. 2011. Are butterflies and moths suitable ecological indicator systems for restoration measures of semi-natural calcareous grassland habitats? Ecological Indicators **11**:1040-1045.

Schulze, P. C., K. J. Wilcox, A. Swift, and J. L. Beckert. 2009. Fast, easy measurements for assessing vital signs of tall grassland. Ecological Indicators **9**:445-454.

Spooner, P. G., and S. V. Briggs. 2008. Woodlands on farms in southern New South Wales: A longer-term assessment of vegetation changes after fencing. Ecological Management & Restoration **9**:33-41.

Stoddard, J. L., D. P. Larsen, C. P. Hawkins, R. K. Johnson, and R. H. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications **16**:1267-1276.

Stol J., Doerr, V. A. J., Davies, M. J., and E. D. Doerr. 2016. Checking for Change: A practical guide to checking whether sites newly managed for conservation are on track to improve. CSIRO, Australia.

Stolton, S., K. H. Redford, and N. Dudley. 2014. The Futures of Privately Protected Areas. IUCN, Gland, Switzerland.

Tongway, D., and N. Hindley. 2004. Landscape function analysis: a system for monitoring rangeland function. African Journal of Range & Forage Science **21**:109-113.

Underwood, E. C., and B. L. Fisher. 2006. The role of ants in conservation monitoring: If, when, and how. Biological Conservation **132**:166-182.

Westgate, M. J., G. E. Likens, and D. B. Lindenmayer. 2013. Adaptive management of biological systems: A review. Biological Conservation **158**:128-139.

Williams, K. J., S. M. Prober, T. D. Harwood, V. A. J. Doerr, T. Jeanneret, G. Manion, and S. Ferrier. 2014. Implications of climate change for biodiversity: a community-level modelling approach. CSIRO Land and Water Canberra.

Zuur, A., E. Ieno, N. Walker, A. Saveliev, and G. Smith. 2009. Mixed Effects Modelling for Nested Data. Pages 101-142 *in* Mixed Effects Models and Extensions in Ecology with R. Springer, New York.

CONTACT US

t 1300 363 400
+61 3 9545 2176
e csiroenquiries@csiro.au
w www.csiro.au

AT CSIRO, WE DO THE EXTRAORDINARY EVERY DAY

We innovate for tomorrow and help improve today – for our customers, all Australians and the world.

Our innovations contribute billions of dollars to the Australian economy every year. As the largest patent holder in the nation, our vast wealth of intellectual property has led to more than 150 spin-off companies.

With more than 5,000 experts and a burning desire to get things done, we are Australia's catalyst for innovation.

CSIRO. WE IMAGINE. WE COLLABORATE. WE INNOVATE.

FOR FURTHER INFORMATION

Land and Water
Veronica Doerr
t +61 2 6246 4099
e veronica.doerr@csiro.au
w www.csiro.au

Land and Water
Jacqui Stol
t +61 2 6246 4058
e jacqui.stol@csiro.au
w www.csiro.au