NZ Sustainability Dashboard Research Report 18/07
Published by ARGOS (Agricultural Research Group on Sustainability)

Effectiveness of farm management actions for enhancing NZ biodiversity: Evidence evaluation assessment

Angela J. Brandt,[1] Catriona J. MacLeod,[1] Lynn V. Dicks[2] and Gorm Shackelford[3]

1. *Manaaki Whenua – Landcare Research, Dunedin, NZ*
2. *University of East Anglia, Norwich, UK*
3. *University of Cambridge, Cambridge, UK*

December 2018

Reviewed by:

Anne-Gaelle Ausseil
Research Priority Leader
Manaaki Whenua – Landcare Research

Approved for release by:

Jon Manhire
NZSD project leader
The Agribusiness Group

Suggested citation for this report

Brandt AJ, MacLeod CJ, Dicks LV, Shackelford G. (2018) Effectiveness of farm management actions for enhancing NZ biodiversity: Evidence evaluation assessment. The NZ Sustainability Dashboard Research Report 18/07. Published by ARGOS.

Acknowledgements

This work was funded by the Ministry of Business, Innovation and Employment (Contract Number AGRB1201).

About this report series

The New Zealand Sustainability Dashboard Report series is a publication of the ARGOS Group (www.argos.org.nz) – as part of the New Zealand Sustainability Dashboard project. All publications can be found on the website nzdashboard.org.nz.

Executive Summary

This pilot study explores how global scientific evidence can be locally adapted for use in New Zealand (NZ). Evidence synthesis is a process for distilling available knowledge in a way that can help decision-makers efficiently find practical solutions to problems, and can be used in multiple ways to support natural resource management practitioners and policy-makers. Here we demonstrate using evidence synthesis in the development of a decision-support tool – a prototype tool for biodiversity assessments on NZ farms being developed as part of the NZ Sustainability Dashboard project.

We quantified the effectiveness of two farm management actions, (1) using shallow or no tillage for cultivation (*Tillage methods*) and (2) the presence of shelterbelts (*Shelterbelts present*), in enhancing biodiversity of five prioritised groups (*Overall biodiversity, Native grassland plants, Native birds of open habitats, Soil life* and *Beneficial insects*). Effectiveness was evaluated using a structured assessment process originally developed by the University of Cambridge for evaluating global Conservation Evidence. We summarised overseas scientific studies of these actions' effects on biodiversity. A panel of six NZ-based biodiversity specialists from a variety of institutions scored this available evidence individually and anonymously in multiple rounds to achieve a consensus of the final classification of an action as more or less likely to be beneficial to a target biodiversity group.

Of the 10 cases of *management action* × *biodiversity group* assessed, seven were categorised as *unknown effectiveness* and three as *trade-offs between benefits and harms* (effect of *Tillage methods* on *Soil life* and *Beneficial insects*, and effect of *Shelterbelts present* on *Beneficial insects*). The reasons for actions' effectiveness being classified as *unknown* was low *certainty* of evidence in one case, low *relevance* of the overseas evidence to NZ in four cases and both low *certainty* and low *relevance* in two cases. Obtaining more evidence from overseas would thus be most informative for the case with low *certainty*, while gathering evidence within NZ would be of greater priority for the remaining six cases.

The results from this pilot study will be used to update the scoring of the prototype biodiversity tool, demonstrating a direct application of evidence synthesis in decision-support tools. Several considerations for future assessments and research programmes have also emerged from this study, including the potential challenges in drawing on overseas evidence databases to evaluate the effects of actions revised to suit the local context. Both the process of compiling evidence and the structured assessment identified key research gaps; for example, the lack of overseas evidence for effectiveness of these two management actions on six of the 10 biodiversity groups prioritised for inclusion in the prototype tool. This work can thus help to prioritise future research programmes with maximum value to NZ practitioners and policy-makers.

Table of Contents

1 Introduction

This pilot study explores how global scientific evidence can be locally adapted for use in New Zealand (NZ). Specifically, it quantifies the effectiveness of two farm management actions in enhancing biodiversity of several target species groups. Studies of these actions' effects on biodiversity from the scientific literature are summarised and provided to a panel of biodiversity specialists. Using a structured assessment process, the available evidence is scored anonymously by each member of the panel. Multiple scoring rounds are undertaken to achieve a consensus of the final classification of an action as more or less likely to be beneficial for a particular biodiversity group. This process was originally developed by the University of Cambridge for evaluating global Conservation Evidence.[1]

1.1 Why evaluate evidence?

Researchers often assume that if they produce evidence, decision makers will find and apply it. However, several behavioural and practical barriers can limit the use of published evidence.[2] Practitioners and policymakers who want to access scientific evidence face the challenge of deriving it from an overwhelmingly large and complex body of available literature. Distilling the available knowledge in a way that helps the decision-makers efficiently find practical solutions to their problem[3] is the first critical step.

Evidence synthesis is the 'process of bringing together information and knowledge from many sources and disciplines to inform debates and decisions'.[4] Systematic reviews and meta-analyses, which are labour-intensive and expensive, provide two mechanisms for addressing this goal. By providing an evidential basis for improvements, such reviews have transformed healthcare[4] and decision-making in social welfare, education, crime and justice and international development.[5]

Unfortunately, some research areas do not naturally lend themselves readily to systematic reviews or meta-analyses, as they have relatively few studies per topic and employ different methods or measure different variables. However, we still need to make sense of the available evidence. The Conservation Evidence project[1] addresses the challenges of sparse and patchy evidence using subject-wide evidence synthesis. The project, based at the University of Cambridge, aims to assess the impact of conservation interventions for all species and habitats worldwide.

Conservation evidence is synthesised in a cost-effective, industrial-scale way (Figure 1), using protocols that combine elements of systematic reviewing and mapping with other techniques, altogether designed to ensure that the evidence produced is inclusive, rigorous, transparent and accessible.[3] Manually searching for relevant studies is costly but still more cost-effective than standard systematic reviews that rely on search terms. However, over time, resources required are reduced as the investigator can build on the efforts of others. Annual or biennial summary updates, for example, are estimated to cost 20% of the initial cost (estimated initial cost of US$70K to 750K for 59 to 457 specific questions addressed).[6]

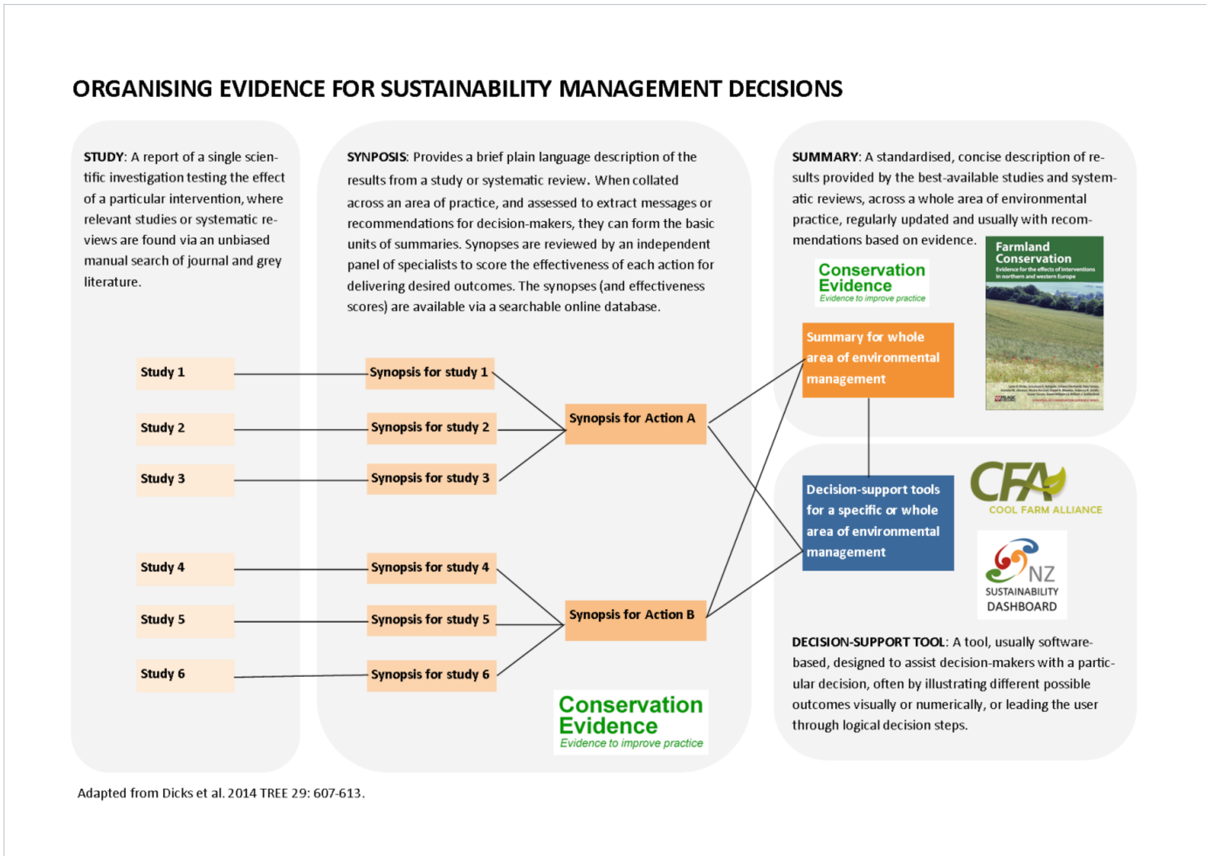


Figure 1: Diagram of process to organise evidence for sustainability management decisions

1.2 Example application of evidence to decision-making

Evidence synthesis can be used in multiple ways to support natural resource management practitioners and policy-makers (Figure 1). Here we explore adapting global evidence for use in the local NZ context with a pilot study involving a decision-support tool currently in development – a prototype tool for biodiversity assessments on NZ farms (Box 1). The prototype tool's scores for implementing a suite of farm management actions were initially derived from results of a structured assessment of the judgement of a panel of biodiversity specialists (Step B1 in Box 1). The results from this pilot evidence synthesis of the effects of two farm management actions (Step B2 in Box 1) will be used to revise the judgement-based scores in the prototype biodiversity tool. A review of the findings from this evidence assessment process and comparison to results from the judgement-based assessment will also highlight areas in which NZ policy and management recommendations might benefit most from further research.

Box 1: Co-designing a biodiversity assessment tool for NZ farms

The *New Zealand Sustainability Dashboard* (NZSD) project is developing a simple, online prototype tool for New Zealand (NZ) farmers to self-assess the management actions they have taken to enhance biodiversity. It also delivers a proof-of-concept of a co-design process for evidence-based tools for NZ farmers and other stakeholders to assess and report their sustainability performance.[7]

Working with a diverse range of NZ stakeholders and researchers, we are capitalising on overseas' research investments to collectively adapt an existing online calculator (the Cool Farm Biodiversity Tool[8] [CFT], developed for north-western European farms) to reflect NZ priorities, farming practices and sectors. This co-design process aims to build trust in the tool and ensure it is widely used. Using the CFT as a standard can provide direct benefits to multiple NZ stakeholders, such as aiding market access and environmental reporting by communicating environmental benefits of farm practices in an industry standard way.

The development of the biodiversity assessment tool consists of three work streams (see diagram below):[9]

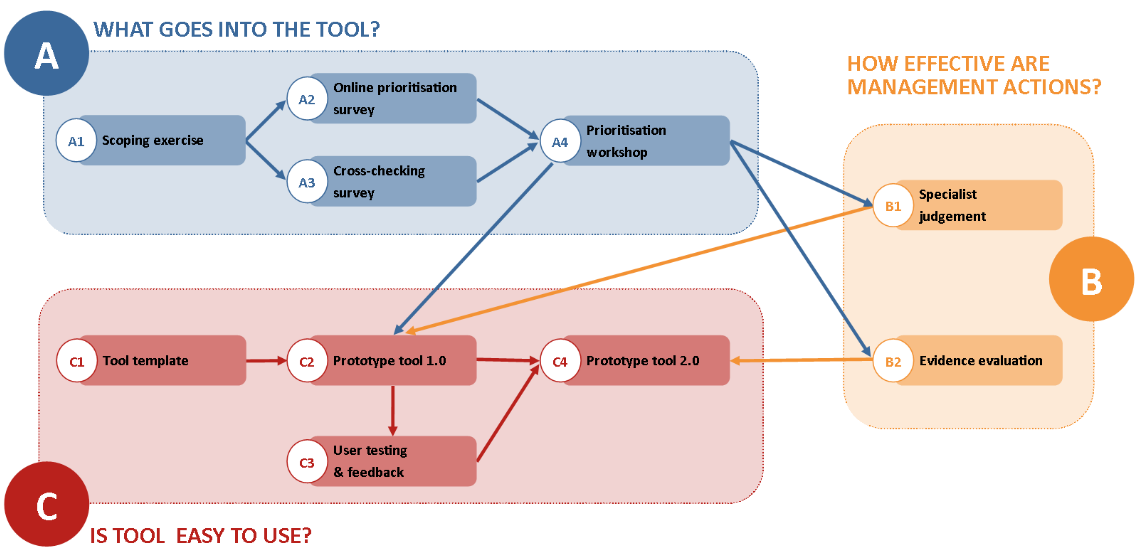
A. *What goes into the tool?* – Tailoring the biodiversity groups and management actions to tell the unique story of NZ's biodiversity. This step, which is complete, involved scoping the possible components (A1) to include in the tool[10] and prioritising components of the prototype tool (A2–A4) with stakeholders.[11]

B. *How effective are management actions?* – Quantifying the expected benefits of a subset of relevant NZ farm management actions for each of the priority biodiversity groups; it involves two substeps:

   B1. A *Specialist Judgement* assessment of the prioritised actions and biodiversity groups,[12] which is complete.

   B2. An *Evidence Evaluation* assessment to determine effectiveness of management actions based on an evaluation of scientific evidence; the focus of this report.

C. *Is the tool easy to use?* – Developing and testing an online prototype tool for biodiversity assessments on NZ farms.



1.3 Why a structured assessment?

Similar to the process of scoping components to include in the prototype biodiversity tool,[10] the benefits of a structured assessment approach include limiting the potential for bias and providing transparency in the tool's development, as well as for any other uses of the evidence synthesis resulting from the assessment. Using a panel of specialists in a formal process of evaluating the effectiveness of farm management actions can minimise the risk of favouring particular management actions that are traditionally recommended or conventionally used but have limited guarantee of success, as well as clarifying how to deal with conflicting evidence for effectiveness of a management action. Our approach involving a specialist panel and a consensus process thus reduces individual bias toward particular management actions and provides a documented framework for the subsequent determination of scores within the prototype tool (i.e. which actions are scored as more or less effective at enhancing a particular biodiversity group).

2 Methods

We adapted the approach of the Cool Farm Biodiversity Tool[13] in this second part of the assessment of effectiveness of farm management actions in enhancing prioritised biodiversity groups for the prototype tool (Step B2 in Box 1). We compiled evidence synopses from actions in the Conservation Evidence database[1] that were closely aligned with a subset of farm management actions assessed in the Specialist Judgement assessment[12] and provided these synopses to a specialist panel to evaluate following a structured assessment process.

2.1 Compiling evidence

We selected two farm management actions from those prioritised by a panel of stakeholder-advisors for inclusion in the prototype biodiversity tool[11] and assessed in the Specialist Judgement assessment[12] – *Tillage methods* and *Shelterbelts present*. These actions were closely aligned to actions in the Conservation Evidence database (Table 1), allowing us to draw on previous research teams' efforts to systematically search the literature and compose synopses of available evidence. These two actions are implemented in different management areas of the farm and are likely to be most relevant to different industry sectors. The Conservation Evidence synopses for these actions included many studies of four taxonomic groups (plants, birds, soil biota and terrestrial invertebrates) that align closely to four of the ecological biodiversity groups prioritised for inclusion in the prototype biodiversity tool (Table 2). The fifth biodiversity group in the assessment (*Overall biodiversity*) includes all taxa potentially occurring within NZ's production landscape, thus the pooled evidence from these four taxonomic groups was considered when evaluating *Overall biodiversity* as a target group.

Table 1: Farm management actions for assessment and aligned actions from Conservation Evidence (CE)

Mgmt area	Index	Action label	Action description	CE category	CE action
Production areas	10	Tillage methods	Use shallow tillage or no tillage as the main method of cultivation	Bee Conservation	Reduce tillage
				Bird Conservation	Reduce tillage
				Farmland Conservation	Reduce tillage
				Mediterranean Farmland	Pest regulation: Use no tillage in arable fields
					Pest regulation: Use reduced tillage in arable fields
					Pest regulation: Use no tillage instead of reduced tillage
					Pollination: Use no tillage in arable fields
					Soil: Use no tillage in arable fields
					Soil: Use reduced tillage in arable fields
					Soil: Use no tillage instead of reduced tillage
				Soil Fertility	Change tillage practices
Small non-production areas	16	Shelterbelts present	Shelterbelts present on farm	Bird Conservation	Plant new hedges
					Plant trees to act as windbreaks
				Farmland Conservation	Plant new hedges
				Mediterranean Farmland	Pest regulation: Plant hedgerows
					Pollination: Plant hedgerows
					Other biodiversity: Plant hedgerows
				Natural Pest Control	Plant new hedges
					Use alley cropping
				Soil Fertility	Plant new hedges
					Use alley cropping

We combined all study summaries from these Conservation Evidence actions in the Bee Conservation,[14] Bird Conservation,[15] Farmland Conservation,[16] Mediterranean Farmland,[17] Natural Pest Control[18] and Soil Fertility[19] synopses. We removed any studies lacking results for biodiversity and removed any results for taxa not included in the target biodiversity groups. We composed synopses for the suite of remaining studies for each action and biodiversity group combination that summarised the studies' geographic locations, main methodological set-ups, direction of effects, response metrics and taxa.

Table 2: Ecological biodiversity groups assessed in the Evidence Evaluation assessment

Ecological group	Description	Example taxa[a]
Native grassland plants	Grasses, flowers, and shrubs native to New Zealand tussock grasslands and open shrublands	Tussocks, herbs, wildflowers, shrubs (e.g. matagouri, *Olearia* spp), harakeke (flax) and other native plant species commonly found in open habitats
Native birds of open habitats	Native birds that mostly use open areas (grasslands or open shrublands) for breeding and feeding	Falcon, harrier hawk, weka, oystercatcher, paradise shelduck, spur-winged plover, gulls
Soil life	Animals, bacteria and fungi that live within the soil, and are mainly found below ground	Earthworms, springtails, mites, fungi, microbes
Beneficial insects	Invertebrates that help agriculture by providing services like pollination or pest control	Bees & other pollinators, spiders, parasitic wasps & other biocontrol agents, ground beetles, millipedes, landhoppers, slaters

2.2 Assembling a specialist panel

To conduct the effectiveness assessment of farm management actions (Step B in Box 1), we compiled a list of scientists from NZ and Australia with expertise in biodiversity and production landscapes who specialised in at least one of the prioritised NZ biodiversity groups that would be assessed. We identified potential candidates for the assessment panel using our familiarity with NZ research networks, recommendations from NZSD project partners and searches of university and other research project websites, as well as further recommendations from specialists we contacted. Of the 23 specialists initially invited, 10 participated in the Specialist Judgement assessment process.[12] We invited this panel and several initial contacts to participate in this Evidence Evaluation assessment, as well as an additional 25 specialists of the four ecological biodiversity groups being assessed (Table 2).

[a] The list of example taxa for each ecological group is not meant to be comprehensive. We recognise that some taxa use multiple habitats; the broad overall habitat preference or requirements of a species should determine its group, but any particular species is not necessarily precluded from belonging to multiple groups.

2.3 Conducting the assessment

We used a method of assessment based on the Delphi technique and used in developing the online Conservation Evidence database, whereby multiple rounds of anonymous individual scoring of management actions are used to achieve a consensus of the final classification of an action as more or less likely to be effective in enhancing biodiversity. The assessment consisted of three rounds of assessor surveys:

1. **Initial scoring.** Assessors scored each management action (from 0 to 100) for overall biodiversity and each ecological group (i.e. n = 10 cases total) to answer the three questions:
    a. How *beneficial* is the practice for the target biodiversity group? *Benefits* could include increased abundance of individuals, enhanced range and diversity of species or occurrence of target species.
    b. How *harmful* is the practice for the target biodiversity group? *Harms* include negative side-effects to the target biodiversity group.
    c. How *certain* is the evidence in the synopsis? *Certainty* was considered a measure of the quality of the evidence provided, including the number of studies, their representativeness (e.g. the variety of agricultural systems and geographic locations, or the methods used) and their robustness (e.g. a replicated, randomised, controlled study with a large sample size might provide higher quality evidence than a non-replicated site comparison), as well as the similarity of results across studies.
    d. How *relevant* is the evidence to the target biodiversity group in New Zealand? When scoring *relevance*, assessors were asked to use their expertise, familiarity with factors supporting or limiting the target biodiversity group and working knowledge of NZ ecology and research to estimate the likelihood that similar studies done in NZ would find the same results as those studies presented in the synopsis.
2. **Agreement with categorisation.** Categories of expected effectiveness using criteria established by Conservation Evidence (Table 3) were assigned to each management action–biodiversity group combination (hereafter, *case*) based on median scores from the first survey round. We used the median rather than the mean to avoid the potential for a skewed mean due to extreme values. Assessors indicated whether they agreed with the category to which each case was assigned.
3. **Final scoring.** Assessors were asked to rescore cases for which there was substantial disagreement with the assigned effectiveness category in the second survey round. The scoring process was identical to round 1; final effectiveness categories were assigned to these cases according to the new median scores as in round 2.

Table 3: Categorisation of farm management actions based on median values of benefits, harms, certainty and relevance scores from the assessment (i.e. on a combination of the size of benefits and harms, and the confidence of assessors in these effects and their relevance to NZ)

Categories	Benefits score	Harms score	Certainty score	Relevance score
Beneficial	≥ 60	< 20	≥ 60	≥ 60
Likely to be beneficial	≥ 60	< 20	40 – 60	≥ 60
OR	40 – 60	< 20	≥ 40	≥ 60
Trade-offs between benefits & harms	≥ 40	≥ 20	≥ 40	≥ 60
Unknown effectiveness	Any score	Any score	< 40	Any score
OR	Any score	Any score	Any score	< 60
Unlikely to be beneficial	< 40	< 20	40 – 60	≥ 60
Likely to be ineffective or harmful	< 40	Any score	≥ 60	≥ 60
OR	< 40	≥ 20	≥ 40	≥ 60

We asked assessors to base their scores solely on the evidence in the synopsis, drawing on their own expertise to interpret that evidence and assess its relevance to NZ taxa. When scoring *benefits* and *harms* of practices on a target biodiversity group, assessors were asked to consider each case independently of other practices or covariates and to consider benefits and harms as independent effects (e.g. where a practice benefits certain species within a target biodiversity group but may harm others). In each survey round, assessors were invited to provide comments giving reasons for their decisions in scoring or agreement/disagreement with assigned effectiveness categories. These comments were provided to all assessors in subsequent survey rounds to help inform their responses.

In survey round 2, there was substantial disagreement with the effectiveness category for one case – the effect of *Tillage methods* on *Soil life*. *Substantial disagreement* was defined as ≥33% of the responding assessors disagreeing with the assigned category. Thus, a third survey round was conducted for this case (indicated in Table 5 in the Appendix) and a new effectiveness category assigned based on the new median scores.

3 Results

3.1 Available evidence

Overall, summaries from 104 studies were included in the synopsis for *Tillage methods* and 24 studies in the synopsis for *Shelterbelts present*, with many individual studies presenting results for multiple ecological biodiversity groups (Figure 2). The studies encompassed a variety of agricultural systems (see Figure 3 for breakdown of 81 *Tillage methods* and 20 *Shelterbelts present* studies with information available in synopses) and occurred in Africa, Asia, Europe, North America and Oceania (see Figure 4 for breakdown of 101 *Tillage methods* and 24 *Shelterbelts present* studies with information available in synopses). However, no studies were from NZ.

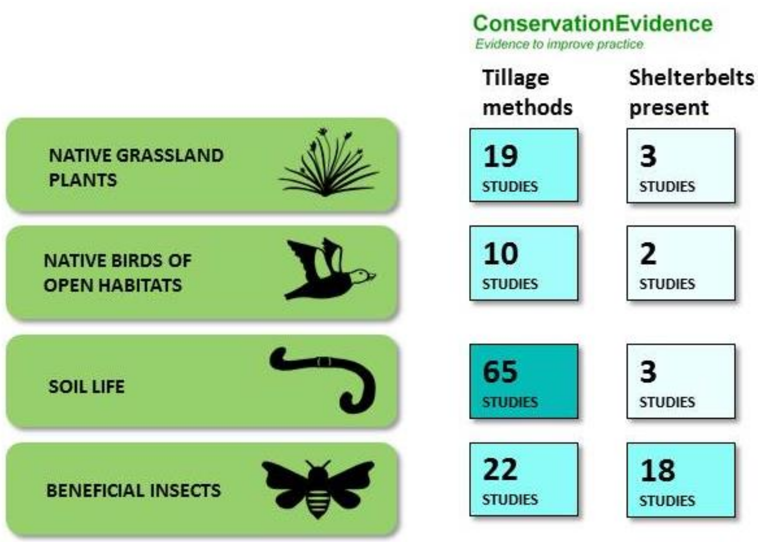


Figure 2: Number of studies available per ecological biodiversity group

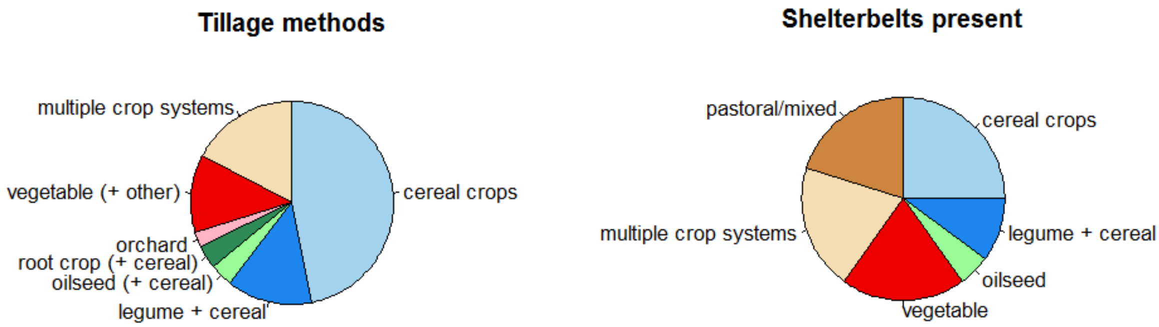


Figure 3: Proportion of studies from different agricultural systems (n = 81 for tillage methods and n = 20 for shelterbelt methods)

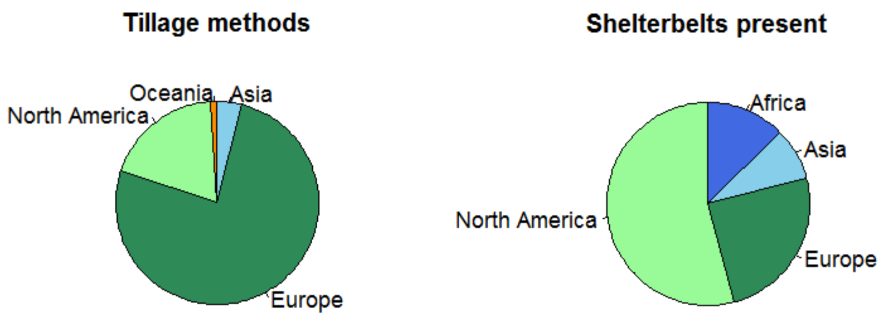


Figure 4: Proportion of studies from different continents (n = 101 for tillage methods and n = 24 for shelterbelt methods)

3.2 Specialist panel composition

The panel consisted of six NZ-based specialists from a variety of institutions (a university, two Crown Research Institutes and a government agency) with expertise in impacts of agricultural practices on NZ biodiversity and specialty in at least one of the biodiversity groups being assessed (Table 4). Two panel members had previously participated in the Specialist Judgement assessment.[12]

Table 4: Specialist panel[b]

Assessor	Affiliation
Dr. Nigel Bell	AgResearch
Assoc. Professor Bruce Burns	University of Auckland
Assoc. Professor Gavin Lear	University of Auckland
Professor William Lee	Manaaki Whenua; University of Auckland
Professor George Perry	University of Auckland
Mr. Bala Tikkisetty	Waikato Regional Council

3.3 Evidence evaluation summary

The complete results of the Evidence Evaluation assessment of effectiveness of two farm management actions at enhancing five biodiversity groups, and relevant Specialist Judgement assessment results for comparison (Table 5). Of the 10 cases (i.e. management actions × biodiversity groups) assessed, seven were categorised as *Unknown effectiveness* and three as *Trade-offs between benefits and harms* (effect of *Tillage methods* on *Soil life* and *Beneficial insects*, and *Shelterbelts present* on *Beneficial insects*).

[b] Note that membership on the assessment panel does not indicate endorsement of the results or the biodiversity tool.

Table 5: Final score summaries from Specialist Judgement and Evidence Evaluation assessments (grey and black text respectively). Asterisks indicate cases rescored in the third assessment round.

Index	Action description	Biodiversity group	Assessment panel scores (median and range)				Effectiveness Category
			Benefits	Harms	Certainty	Relevance	
10	Use shallow tillage or no tillage as the main method of cultivation	Overall biodiversity	30 (10 - 100)	0 (0 - 90)	30 (5 - 90)	–	Unknown effectiveness
			65 (60 - 90)	27.5 (15 - 60)	65 (40 - 90)	55 (20 - 80)	Unknown effectiveness
		Native grassland plants	10 (0 - 30)	0 (0 - 80)	50 (5 - 80)	–	Unlikely to be beneficial
			52.5 (40 - 65)	30 (15 - 50)	50 (20 - 70)	42.5 (10 - 80)	Unknown effectiveness
		Native birds of open habitats	25 (10 - 60)	15 (0 - 80)	40 (5 - 90)	–	Unlikely to be beneficial
			62.5 (40 - 80)	10 (0 - 30)	60 (20 - 60)	35 (10 - 70)	Unknown effectiveness
		Soil life	60 (50 - 100)	0 (0 - 50)	70 (30 - 90)	–	Beneficial*
			70 (60 - 80)	25 (10 - 40)	72.5 (50 - 80)	72.5 (50 - 80)	Trade-offs between benefits and harms*
		Beneficial insects	25 (10 - 80)	15 (0 - 60)	35 (10 - 90)	–	Unknown effectiveness
			60 (20 - 70)	35 (15 - 60)	65 (35 - 80)	62.5 (40 - 80)	Trade-offs between benefits and harms
16	Shelterbelts present on farm	Overall biodiversity	30 (5 - 100)	5 (0 - 30)	50 (30 - 100)	–	Unlikely to be beneficial
			55 (50 - 95)	20 (0 - 30)	50 (20 - 100)	45 (10 - 100)	Unknown effectiveness
		Native grassland plants	0 (0 - 20)	0 (0 - 20)	50 (10 - 70)	–	Unlikely to be beneficial
			50 (30 - 80)	22.5 (0 - 40)	25 (20 - 90)	25 (5 - 90)	Unknown effectiveness
		Native birds of open habitats	25 (0 - 100)	7.5 (0 - 40)	50 (10 - 100)	–	Unlikely to be beneficial
			50 (50 - 80)	50 (10 - 50)	20 (10 - 80)	35 (10 - 60)	Unknown effectiveness
		Soil life	50 (10 - 100)	0 (0 - 50)	40 (10 - 100)	–	Likely to be beneficial
			50 (15 - 70)	0 (0 - 40)	35 (20 - 100)	60 (20 - 100)	Unknown effectiveness
		Beneficial insects	50 (5 - 100)	0 (0 - 20)	65 (10 - 100)	–	Likely to be beneficial
			65 (50 - 70)	20 (0 - 50)	70 (40 - 90)	62.5 (40 - 80)	Trade-offs between benefits and harms

3.4 Actions with unknown effectiveness

Two criteria could lead to a farm management action being categorised as *Unknown effectiveness* – low quality of evidence (i.e. median *certainty* score < 40) and/or low relevance to target biodiversity group in NZ (i.e. median *relevance* score < 60). Of the seven cases categorised as *Unknown effectiveness*, one was assessed as low in certainty, four as low in relevance and two as low in both certainty and relevance (Table 6). These distinctions are informative about whether obtaining more evidence from overseas will be sufficient to estimate whether these actions are expected to enhance NZ biodiversity in these target groups. Obtaining evidence within NZ is most needed for the cases with low relevance scores, because even where there is strong evidence for benefits or harms in the overseas literature, its low relevance to the target biodiversity group renders it less useful within NZ. It is possible that cases with both low certainty and low relevance scores could be better informed by obtaining more evidence from overseas (e.g., because the low certainty may derive from low geographic representation or mixed results among studies). However, efforts to obtain more overseas evidence would be best focused on the case with a low certainty score but high relevance to NZ, as additional evidence for this case is most likely to result in a revised categorisation of effectiveness.

Table 6: Reasons for categorisation of "Unknown effectiveness"

Farm management action	**Biodiversity group**	**Scores below threshold**	
		Certainty	**Relevance**
Use shallow tillage or no tillage as the main method of cultivation	Overall biodiversity		✓
	Native grassland plants		✓
	Native birds of open habitats		✓
Shelterbelts present on farm	Overall biodiversity		✓
	Native grassland plants	✓	✓
	Native birds of open habitats	✓	✓
	Soil life	✓	

3.5 Comparisons to Specialist Judgement assessment

Only one case of the subset assessed in the Evidence Evaluation assessment was categorised with the same expectation of effectiveness in the two assessment approaches – effect of *Tillage methods* on *Overall biodiversity* (Table 7). Cases categorised as *Unlikely to be beneficial* in the Specialist Judgement assessment were all categorised with 'Unknown effectiveness' in the Evidence Evaluation assessment, primarily because the overseas evidence available was deemed of low relevance to the NZ context (see Table 6).

The Evidence Evaluation assessment produced higher harms scores for two cases than the Specialist Judgement assessment (see Table 5), thus the expected beneficial effects of *Tillage methods* on *Soil life* and *Shelterbelts present* on *Beneficial insects* appear to be reduced when considering available evidence. The Evidence Evaluation assessment may also have clarified the effects of *Tillage methods* on *Beneficial insects*, which was categorised with *Unknown effectiveness* in the Specialist Judgement assessment.

This comparison further highlights that gathering more evidence to determine the effects of *Shelterbelts present* on *Soil life* should be of high priority – according to Specialist Judgement, this action is expected to have benefits in NZ and the evidence available from overseas has high relevance but is currently not sufficient to categorise the action as beneficial.

4 Considerations for future assessments

As with the Specialist Judgement assessment, this pilot Evidence Evaluation assessment process has suggested several considerations for future assessments. Comments from the assessment panel were particularly helpful to inform these reflections.

4.1 Consequences of adapting actions to local context

An essential aspect of adapting the CFT to the NZ context was to revise the farm management actions considered for the tool to match local priorities – to tell the NZ biodiversity story. A consequence of revising these actions, however, was that many actions in the prioritised list[11] no longer closely align with actions in the Conservation Evidence database.[1] This limits the number of actions for which local researchers can draw on the study summaries provided in Conservation Evidence to conduct a new assessment. Thus, new searches of the literature and construction of study summaries would be required to conduct an evidence evaluation for the complete prioritised list of farm management actions that have gone into the NZ biodiversity assessment tool.

Table 7: Effectiveness of actions in Specialist Judgement vs. Evidence Evaluation assessments

Action description	Biodiversity group	Effectiveness Category	
		Specialist Judgement	Evidence Evaluation
Use shallow tillage or no tillage as the main method of cultivation	Overall biodiversity	Unknown effectiveness	Unknown effectiveness
	Native grassland plants	Unlikely to be beneficial	Unknown effectiveness
	Native birds of open habitats	Unlikely to be beneficial	Unknown effectiveness
	Soil life	Beneficial	Trade-offs between benefits and harms
	Beneficial insects	Unknown effectiveness	Trade-offs between benefits and harms
Shelterbelts present on farm	Overall biodiversity	Unlikely to be beneficial	Unknown effectiveness
	Native grassland plants	Unlikely to be beneficial	Unknown effectiveness
	Native birds of open habitats	Unlikely to be beneficial	Unknown effectiveness
	Soil life	Likely to be beneficial	Unknown effectiveness
	Beneficial insects	Likely to be beneficial	Trade-offs between benefits and harms

4.2 Categorisation of effectiveness

The effectiveness of each farm management action in enhancing biodiversity of the target group was placed into one of six categories based on score thresholds used by Conservation Evidence (Table 3), which have been adapted from the Clinical Evidence Handbook.[20] The intention of the categories is to distinguish the potential net benefits of a management action (*beneficial* vs. *not beneficial/ineffective*) from the certainty of assessors in the outcome (*likely* vs. *unlikely*). The category thresholds of 20, 40 and 60 were established in previous expert assessments conducted by Conservation Evidence researchers and were used in the Specialist Judgement and Evidence Evaluation assessments to provide continuity with the Conservation Evidence approach and the CFT Biodiversity module.

Future assessments and subsequent uses of the current assessment results could apply different score thresholds based on acceptability by stakeholders. For example, if a 30% likelihood of harm occurring to a target biodiversity group is acceptable when the likelihood of benefit is greater than 60%, then both *Tillage methods* on *Soil life* and *Shelterbelts present* on *Beneficial insects* would be categorised as *Beneficial*.

4.3 Evaluating evidence from synopses

Multiple assessors provided feedback that it was difficult to systematically and accurately assess the evidence given the format of synopses and study summaries provided – at least two specialists (one who participated and one who declined to participate) suggested that a more quantitative presentation of the evidence, such as a meta-analysis, would be helpful. A strength of the current approach, however, is that it can use the full suite of studies available, with different methods and measured responses, to inform the assessment;[3] using meta-analyses alone would likely limit the scope of the assessment because they require standardised methods and response variables, as well as a large sample size. Future work could explore how to incorporate both types of approaches into an evidence assessment to draw on the strengths of each.

4.4 Relevance of overseas research to NZ context

As noted in Table 6, the primary reason for categorising the effectiveness of farm management actions as *Unknown* is low *relevance* (according to assessor scores) of overseas evidence to the NZ context. There are several factors that may determine whether study outcomes could be expected to be similar in different locations (e.g. similar climates, agricultural systems or species ecology), and the Conservation Evidence team is currently exploring ways to assess relevance of study synopses to a new local context. Assessors in this Evidence Evaluation assessment scored high *relevance* when overseas studies occurred in similar climates to NZ and the ecology of the target biodiversity group in NZ was expected to be similar to that from overseas (e.g. earthworms, soil microbes). Conversely, assessors scored low *relevance* when ecology of the target biodiversity group in NZ was unlikely to match that from overseas (e.g. ruderals are rare among native grassland plants, native birds of open habitats in NZ are generally in different trophic groups than in Europe). Thus, future assessments of global evidence for actions implemented in a local context could be improved by including more detail about environmental context and ecology of target biodiversity groups, and by exploring the different aspects of *relevance* in more detail and for each individual study.

4.5 Identifying research gaps

Both the process of compiling evidence and the structured assessment can identify key research gaps, which could be prioritised for future research to maximise value to NZ practitioners and policy-makers. The overseas evidence available for the two farm management actions assessed here lacked any studies from NZ and examined effects on only four of the 10 ecological biodiversity groups prioritised for inclusion in the prototype tool,[11] with limited evidence of the action *Shelterbelts present* for three of these four groups. The assessment panel scored overseas evidence as being of low *relevance* in six of the 10 cases assessed (Table 6). These results thus indicate where to focus searches of NZ-based evidence and funding of new studies on biodiversity in production landscapes to achieve the greatest contribution to decision-making. A systematic search of NZ-based evidence also has the potential to fill gaps in the global evidence base (e.g. the Conservation Evidence database), such as in effects of shelterbelts on diversity of different bird groups.[21]

5 Next steps

The direct application of results from this pilot study and the future research directions that it suggests have potential for great value to a variety of end-users, such as central and local government, land managers in government agencies, non-governmental organisations, industry bodies, farmers and growers, researchers, funding bodies and community groups.

5.1 Informing the prototype biodiversity tool

The results from this Evidence Evaluation assessment will be used to update the scoring of the online prototype biodiversity assessment tool.[c] The effectiveness categories to which each farm management action was assigned for each biodiversity group in the Specialist Judgement assessment[10] were translated into a score of expected benefit for that biodiversity group to develop the tool. Scores from the cases assessed in the Evidence Evaluation assessment can be added to these baseline scores to give greater credit to actions whose benefit is supported by scientific evidence, as was done in the CFT Biodiversity module.[9]

[c] https://landcare.shinyapps.io/BiodivPrototype/

5.2 Addressing research gaps

The lessons learnt from the Evidence Evaluation assessment will be used to inform future assessments and research programmes. Future assessments might include systematic searches of both overseas and NZ evidence for effects of farm management actions in the biodiversity tool that do not closely align with actions in the Conservation Evidence database. Future research programmes might then focus on conducting studies that can fill gaps in available evidence, particularly within NZ.

6 References

[1] www.conservationevidence.com

[2] Sutherland WJ, Wordley CFR. (2017) Evidence complacency hampers conservation. *Nature Ecology & Evolution* 1: 1215-1216.

[3] Sutherland WJ, Wordley CFR. (2018) A fresh approach to evidence synthesis. *Nature* 558: 364-366.

[4] Donnelly CA, Boyd I, Campbell P, Craig C, Vallance P, Walport M, Whitty CJM, Woods E, Wormald C. (2018) Four principles to make evidence synthesis more useful for policy. *Nature* 558: 361-364.

[5] https://campbellcollaboration.org/better-evidence.html

[6] Dicks LV, Walsh JC, Sutherland WJ. (2014) Organising evidence for environmental management decisions: a '4S' hierarchy. *Trends in Ecology and Evolution* 29: 607-613.

[7] Brandt A, MacLeod C, Collins K. (2017) Developing a simple on-farm biodiversity assessment tool. NZ Sustainability Dashboard Research Summary 17/13. ARGOS. 4 p. http://www.nzdashboard.org.nz/uploads/2/3/7/3/23730248/biodiversity_tool_background.pdf

[8] https://coolfarmtool.org/coolfarmtool/biodiversity/

[9] Brandt A, MacLeod C, Collins K. (2017) Process to co-design an evidence-based biodiversity assessment tool. NZ Sustainability Dashboard Research Summary 17/14. ARGOS. 5 p. http://www.nzdashboard.org.nz/uploads/2/3/7/3/23730248/biodiversity_tool_methods_summary.pdf

[10] Brandt AJ, MacLeod CJ. (2018) Telling the New Zealand farmland biodiversity story: Scoping relevant components for an online assessment tool. The NZ Sustainability Dashboard Research Report 18/02. Published by ARGOS. https://datastore.landcareresearch.co.nz/dataset/report-scoping-biodiversity-components

[11] MacLeod CJ, Brandt AJ, Collins K, Dicks LV. (2018) Telling the New Zealand biodiversity story: Prioritisation workshop for what goes into the online assessment tool. NZ Sustainability Dashboard Research Report 18/04. Published by ARGOS. https://datastore.landcareresearch.co.nz/dataset/biodiversity-prioritisation-workshop

[12] Brandt AJ, MacLeod CJ, Dicks LV, Shackelford G. (2018) Effectiveness of farm management actions for enhancing NZ biodiversity: Specialist judgement assessment. The NZ Sustainability Dashboard Research Report 18/06. Published by ARGOS. https://datastore.landcareresearch.co.nz/dataset/effectiveness-of-management-actions

[13] Dicks et al., unpublished

[14] Dicks LV, Showler, DA, Sutherland WJ. (2010) *Bee conservation: evidence for the effects of interventions*. Pelagic Publishing, Exeter, UK.

[15] Williams DR, Pople RG, Showler DA, Dicks LV, Child MF, zu Ermgassen EKHJ, Sutherland WJ. (2012) *Bird Conservation: Global evidence for the effects of interventions*. Exeter, Pelagic Publishing.

[16] Dicks LV, Ashpole JE, Dänhardt J, James K, Jönsson A, Randall N, Showler DA, Smith RK, Turpie S, Williams D, Sutherland WJ. (2013) *Farmland Conservation: Evidence for the effects of interventions in northern and western Europe*. Exeter, Pelagic Publishing.

[17] Shackelford GE, Kelsey R, Robertson RJ, Williams DR, Dicks LV. (2017) *Sustainable Agriculture in California and Mediterranean Climates: Evidence for the effects of selected interventions*. Synopses of Conservation Evidence Series. University of Cambridge, Cambridge, UK.

[18] Wright HL, Ashpole JE, Dicks LV, Hutchison J, McCormack CG, Sutherland WJ. (2018) Some Aspects of Enhancing Natural Pest Control. Pages 559-582 in: W.J. Sutherland, L.V. Dicks, N. Ockendon, S.O. Petrovan & R.K. Smith (eds) *What Works in Conservation 2018*. Open Book Publishers, Cambridge, UK.

[19] Key G, Whitfield M, Dicks LV, Sutherland WJ, Bardgett RD. (2017) Enhancing Soil Fertility. Pages 383-404 in: W. J. Sutherland, L. V. Dicks, N. Ockendon & R. K. Smith (eds) *What Works in Conservation 2017*. Open Book Publishers, Cambridge, UK.

[20] *Clinical Evidence Handbook*. BMJ Publishing Group: London, UK, 2013.

[21] MacLeod CJ, Blackwell G, Benge J. (2012) Reduced pesticide toxicity and increased woody vegetation cover account for enhanced native bird densities in organic orchards. *Journal of Applied Ecology* 49: 652-660.