special issue:
Advances in tools for bird population monitoring in New Zealand

Designing a bird monitoring scheme for New Zealand's agricultural sectors

Catriona MacLeod[1,2,*], Grant Blackwell[1,3], Florian Weller[1] and Henrik Moller[1]

[1] Agriculture Research Group on Sustainability, Centre for Study of Agriculture, Food and Environment, University of Otago, PO Box 56, Dunedin 9054, New Zealand
[2] Landcare Research, Private Bag 1930, Dunedin 9054, New Zealand
[3] Office of the Parliamentary Commissioner for the Environment, PO Box 10-241, Wellington 6143, New Zealand
*Author for correspondence (Email: macleodc@landcareresearch.co.nz)

Published on-line: 30 July 2012

Abstract: Growing concerns about significant biodiversity decline due to agricultural intensification are increasingly leading consumers to seek agricultural products that are produced sustainably. To raise awareness of sustainable land management and direct policy and research to mitigate adverse impacts, large-scale bird monitoring programmes are being used in Europe. New Zealand's first farmland bird monitoring scheme was established in 2004 to quantify bird abundance on 98 farms across three sectors (sheep & beef, dairy and kiwifruit). Distance methods were considered ideal because they minimised disruption by nuisance variables that affected detectability (most often observer and whether birds were seen or heard; less frequently, effects of wind, habitat and farming systems). However, distance detection functions could only be measured for half the species present on the study farms, and sampling uncertainty remained high for several of those species. Gradually more species with reduced sampling uncertainty can be added as sufficient detections are gathered to generate their global detection functions. This will likely increase the scheme's power to detect any ongoing decline, but simulations that combine sampling uncertainty with observed inter-annual variation in abundance are now needed to test whether population-decline thresholds can be reliably detected using the current and alternative survey designs.

Keywords: detectability; distance sampling; effective survey width; sampling uncertainty; spatial scale

Introduction

There is growing concern globally about the adverse effects of agricultural intensification on biodiversity and ecosystem services (e.g. Tilman 1999; Foley et al. 2005; Butler et al. 2007). In Europe, for example, agricultural intensification has been a major driver of population declines of a wide range of bird, invertebrate and plant species (e.g. Krebs et al. 1999; Chamberlain et al. 2000; Sotherton & Self 2000; Donald et al. 2001; Benton et al. 2002). Such concerns are increasingly leading consumers to seek agricultural products grown using sustainable land management practices, making development of suitable socio-environmental systems and indicators (e.g. ecolabelling) that inform on product sustainability a priority (e.g. Golden et al. 2010).

Birds are a potential focal group or species for environmental monitoring programmes (Furness & Greenwood 1993), because (1) they are good indicators of wider ecosystem health and functioning; (2) they are generally well recognised and familiar to farmers, politicians and the public; and (3) some species have potential as indicators of good farming system practices for increased market access for farm produce. In the UK, the wild bird index is one of 15 headline indicators of 'quality of life' recently introduced by the government for measuring the country's progress towards sustainable development (DEFRA 2002). By simplifying large amounts of scientific data into a simple, understandable and meaningful index, the bird indicator was initially intended as a tool to raise awareness of sustainable land management issues (Gregory et al. 2004; http://ww.defra.gov.uk). Having successfully engaged media and public interest, the indicator has since been used to set research and management targets as well as monitor progress. Similar pan-European indicators for bird populations have also been developed to inform both management and policy in the European Union at various regional scales (Gregory et al. 2005; http://www.ebcc.info).

The UK's wild bird index is based on information collected since the 1970s by skilled volunteers, as part of an integrated national population monitoring scheme (Baillie 1990). These data have shown large-scale declines in the geographic range and population size of lowland farmland bird species, with similar patterns and timing of declines observed across a wide range of species (e.g. Gibbons et al. 1993; Fuller et al. 1995; Siriwardena et al. 1998; Fewster et al. 2000). Intensive

This special issue reviews the advances in tools for bird population monitoring in New Zealand. This issue is available at www.newzealandecology.org.nz/nzje/.

autecological studies, integrated modelling of population surveys, a nest recording scheme and ring scheme data have all demonstrated that agricultural intensification has reduced survival and/or reproductive success. For example, increased annual mortality, linked to the loss of seed-rich habitats such as over-wintered stubble, has been identified as the key demographic mechanism behind the significant and widespread decline of many farmland bird species (Siriwardena et al. 2000). Recent studies have shown that increasing the availability of winter food supplies can increase survival and local abundance (Peach et al. 2001; Hole et al. 2002) and, to some extent, can influence trends in breeding populations (Siriwardena et al. 2007).

In New Zealand, where production lands account for 58% of the area, recent studies have identified an ongoing and accelerating trend for agricultural intensification (e.g. PCE 2004; MacLeod & Moller 2006). However, despite various calls for the development of a monitoring scheme that provides reliable biodiversity and environmental indicators of the impact of land use changes on native and introduced taxa (Meurk & Swaffield 2000; Norton & Miller 2000; Perley et al. 2001; PCE 2004, 2010; Moller et al. 2005, 2008), neither the nature of this threat nor the extent of its impact on biodiversity is known (MacLeod et al. 2008; Moller et al. 2008). Most avian research effort to date has focused on threatened species within the conservation estate (MacLeod et al. 2008), so there is very little or no information available on the population trends of bird species associated with farmland habitats at either regional or national scales (MacLeod et al. 2011). Knowledge of status, composition or size of bird populations in farmland areas, or the factors impacting them, is also lacking. Hence, it is currently difficult to identify which land practices are sustainable and which species could be used as sustainability indicators.

As part of a broader research programme examining the environmental, social and economic sustainability of New Zealand's farming systems, the Agricultural Research Group on Sustainability (ARGOS) initiated a farmland bird monitoring scheme in 2004. This scheme aims initially to establish baseline information on community composition and species distribution and abundance in relation to different farming systems and locations, as well as other habitats and countries (e.g. MacLeod et al. 2009). The longer-term goals are to determine the drivers of variation in bird abundance and diversity (MacLeod et al. 2012), to identify focal species that can be used as indicators for monitoring the impact of land use change, and to see how these can be integrated with similar economic and social indicators to understand drivers of change.

Here we outline the design of the ARGOS bird monitoring scheme and the distance-sampling protocols used to assess bird abundance on farms in three sectors (kiwifruit, sheep & beef and dairy) in the field. We then outline the data analysis methods used to account for variation in detection probabilities (among species, surveys and sectors) and extract density estimates. The pros and cons of varying the spatial and temporal resolution of the data included in these analyses are also explored. A parallel study investigated the need for distance-sampling protocols to account for heterogeneity in detectability (Weller 2012; Weller et al. 2012), so the trade-offs between raw counts versus density estimates are not considered here.

Methods

Study sites

Bird surveys were undertaken on 98 properties from three different agricultural sectors (37 sheep & beef farms [SB], 24 dairy farms [DY] and 37 kiwifruit orchards [KF]; Table 1). The geographical distribution of study sites throughout New Zealand was broadly representative of land-use patterns at the national scale, with all sheep & beef farms located on the South Island and all dairy farms and most kiwifruit orchards on the North Island (Fig. 1). Within each sector, the study sites were located within 12 clusters. Each cluster consisted of 2–4 properties, with each property managed under a different regime ('conventional', 'integrated' or 'organic') or undergoing a conversion from a conventional to an organic regime (Table 1). Properties within the same cluster were matched, as closely as possible, according to location, soil type, topography and climate. However, properties were not matched according to any other management practices (e.g. habitat composition), to avoid the risk of excluding the very differences that may be responsible for creating the observed variability in biodiversity in the first place (Unwin et al. 1995, cited in Hole et al. 2005).

Each sheep & beef cluster contained one conventional, one integrated management and one organic farm, while each dairy cluster included one conventional and one organic conversion farm (Table 1). Each kiwifruit cluster had an integrated management 'Hayward' (Green; *Actinidia deliciosa*), an integrated management 'Hort 16A' (Gold; *Actinidia chinensis*) and an organically managed 'Hayward' orchard (see Carey et al. (2009) for detailed description of management systems). It is important to note that half of the Gold orchards also grew some Green kiwifruit vines within the same area we surveyed for birds. Among the Gold orchards, on average 62% of the area planted in vines was of the Gold variety (but ranged from 21% to 100%). Property sizes varied between sectors, with sheep & beef properties being the largest and kiwifruit ones the smallest (Table 1).

Bird surveys

Over a 6-year period (2004–2010), bird surveys were undertaken once on the dairy farms and three times on the sheep & beef farms and the kiwifruit orchards (Table 1). Bird surveys were carried out along line transects, using a distance-sampling technique (Buckland et al. 2001), during the breeding season (November–February). This survey effort focused primarily on assessing bird community composition within the production areas of each property.

In kiwifruit orchards, the first transect line was initiated at a random point within 50 m of one of the property's boundary corners. The majority of transects ran parallel to the kiwifruit vines, starting and ending at the property boundary. Subsequent transects were located 50 m apart and ran parallel to the preceding ones. Where time constraints prevented complete coverage of all blocks on a property, priority was given to kiwifruit blocks that were of the same classification as the orchard (i.e. Green, Gold or Organic), then blocks of other kiwifruit types (i.e. Green blocks that were part of a Gold orchard), and finally blocks of other orchard crops (e.g. avocado, citrus). To maximise the number of detections per orchard, repeat surveys within the same season were undertaken on each property along the same transects in 2004/05 and different transects in 2006/07 (with the latter using a new random start-point for each survey).

Table 1. Survey effort summary, specifying number and distribution of study sites, total number of transects and transect length surveyed per survey within each sector, and number of transects and total transect length per property per survey within each sector of New Zealand agriculture. ([a]Note: in the initial survey on the kiwifruit orchards, the same transects were resurveyed up to three times (each time by a different observer); the total transect length per property is equal to the product of the transect length and the frequency of survey events.)

Variable	Measure	Dairy	Kiwifruit			Sheep & beef		
		2006/07	2004/05	2006/07	2009/10	2004/05	2007/08	2009/10
Survey dates		9 Jan. –3 Feb.	16 Nov. –21 Jan.	28 Nov. –21 Jan.	6–15 Nov.	17 Nov. –30 Jan.	3 Dec. –27 Jan.	11 Nov. –18 Jan.
Clusters (*N*)		12	12	12	12	12	12	11
Properties (*N*)	Total	24	37	37	36	37	34	26
	Within panel	12 Conventional	12 Gold	12 Gold	12 Gold	12 Conventional	11 Conventional	9 Conventional
			12 Green	12 Green	12 Green	12 Organic	11 Organic	8 Organic
			12 Organic	12 Organic	12 Organic	12 Integrated	11 Integrated	8 Integrated
		12 Converting to organic	1 Converting to organic	1 Converting to organic		1 Converting to organic	1 Converting to organic	1 Converting to organic
	Median per cluster	2	3	3	3	3	3	2
Property area (ha)	Median	102	6.5	6.5	6.5	393	405	447
	Minimum	40	1.4	1.4	1.4	141	141	141
	Maximum	590	24.9	24.9	24.9	1631	1631	1631
Transects (*N*)	Total across sector	225	228[a]	353	246	320	372	239
	Median per property	9	6[a]	10	6	9	11.5	9
	Minimum per property	6	2	2	1	5	8	3
	Maximum per property	12	13	15	27	11	16	15
Transect length (m)	All properties	49 730	69 482[a]	54 256	36 928	151 455	183 681	113 130
	Median per property	2120	1733	1370	844	4000	5681	4406
	Minimum per property	1346	507	201	325	2260	4000	1345
	Maximum per property	2810	4848	3484	3822	5500	8000	7123
Observers (*N*)	Total	4	4	5	5	8	4	5
	Median per cluster	4	4	4	3	4	3	5
	Median per property	3	3	1	1	4	3	4

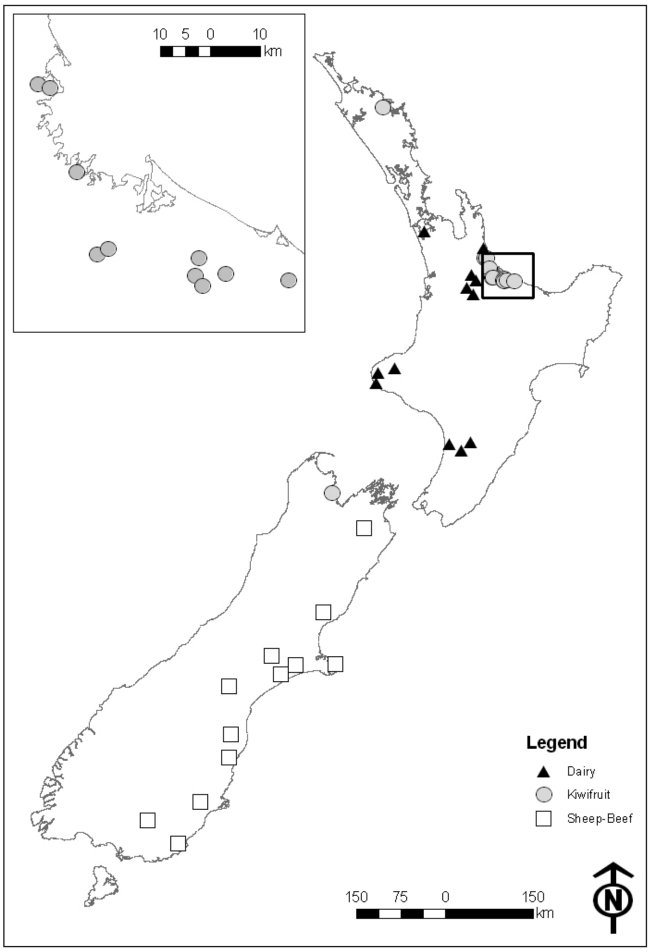

Figure 1. Study sites within each sector (sheep & beef, dairy, kiwifruit) in New Zealand were located within 12 clusters, each consisting of 2–4 properties managed under a different regime (conventional, integrated, organic, or converting to organic; see Table 1).

Sampling effort for the preliminary surveys described here was set mainly by pragmatic and resource-limit considerations – the budget enabled us to station counting teams of three to five observers at each cluster, with the number of observers surveying an individual property within each cluster varying among sectors and surveys (Table 1). On the sheep & beef and dairy farms, observers were asked to survey 10–15 transects per property, with survey effort varying roughly in proportion to the property area (Table 1). A new set of transect start-points were randomly generated for each survey, with points on each property separated by ≥ 200 m and transects ≥ 100 m from the farm boundary. Observers navigated to each start-point (using Garmin eTrex GPS) and then walked a 100–500 m transect due south or north (Table 1). Transect start-points, orientation and lengths were determined by accessibility and the property width. Points that were rejected due to physical obstacles or health and safety concerns were replaced with other randomly selected points.

Surveys were carried out between 0800 and 1600 hours by one or more trained observers to maximise the time available for conducting the surveys. The peak calling periods at dawn and dusk, when conspicuousness and detectability can change rapidly, were avoided (Dawson & Bull 1975). Properties within each cluster were surveyed either concurrently or on subsequent days, with observers rotated between properties and management systems to control for potential observer bias. On all sites, the current ambient temperature, average wind speed (km hr^{-1}) and relative humidity were recorded using a Kestral 4000 portable weather meter (Nielsen-Kellerman, PA). Relative cloud cover (on a six-point scale where 0 = no cloud and 5 = complete cloud cover) and weather conditions (fine, overcast, raining) were also recorded.

For each independent detection of an individual bird or flock of birds, the observer recorded the detection cue (seen or heard or both), the number of individuals (flock size), and the behaviour and location of the bird. Observers used a range finder (Bushnell Yardage Pro®, Bushnell Performance Optics, Overland Park, KS, USA) to record the angle of the bird from the transect line and the distance to the location where the bird (or the centre point of a flock of birds) was first detected. For birds that were heard singing or calling (but not seen) within a clearly defined habitat feature, the distance to that feature was measured. No distance measures were recorded when the observer was uncertain about the location of the bird(s). Any observations of birds flying overhead (i.e. not associated with any specific feature on the property) were recorded but then excluded from the dataset prior to analyses. The distance and angle data were used to calculate the perpendicular distance of the bird from the transect line. In the kiwifruit orchards, the data recording process was simplified after the initial survey, with the observer only required to note the location of the bird (relative to the transect) within six distance bands: 0–5, 6–15, 16–25, 26–50, 51–100 and > 100 m. Bird observations from the initial kiwifruit survey were also subdivided into these distance bands for data analysis. To minimise the risk of sampling bias associated with only deleting records with missing values, any transects with several incomplete detection records were removed from the dataset prior to analyses.

Constructing bird detection functions

Raw counts may be informative for measuring trends, provided there are no systematic biases in detection probabilities over time. However, using raw counts to calculate bird densities usually results in biased estimates because this approach assumes that detectability is constant among species and habitat types (Buckland et al. 2001, Norvell et al. 2003). On the ARGOS farms, for example, Weller et al. (2012) show heterogeneity in detectability in relation to habitat composition for a subset of species. Thus, to calculate more robust measures of density in this study, distance-sampling software (Distance version 6.0; Thomas et al. 2010) was used to model variation in detection probabilities among species and habitats. More specifically, the decline in detectability of individuals or flocks of individuals with distance from the transect line was modelled by fitting detection functions to the distribution of bird detections data.

A minimum of 40–60 detections is recommended when fitting a detection function for species sampled using distance-sampling transects (Buckland et al. 2001). For each species (excluding domestic and feral ones) with sufficient observations, we fitted: (1) a survey-specific detection function for each sector and survey independently, to allow for subtle differences in sampling method over time associated with changes in field team composition; (2) a global detection function for each sector independently, to test whether increasing the number of detections per species (by pooling data from all surveys) resulted in more precise density estimates and increased the number of species that could be modelled; (3) global and survey-specific detection functions to the data collected from the 26 focal sheep & beef farms (26SBs) that were surveyed throughout the study, to assess the impact of the decline in the number of study sites (from 37 to 26) in this sector over the 6-year period.

To optimise model fit, bird observations from the sheep & beef and dairy sectors were truncated (using a minimum detection probability threshold of 0.15; Buckland et al. 2001; Appendices 1–4) and subdivided into distance intervals of varying frequency (3–10) and dimensions. Similarly, for the kiwifruit sector, bird detections >100 m from the transect were excluded from the analyses and two or more distance bands were often combined to aid model fit.

A three-step process was used to identify the ‘best-fit’ detection function for each species. First the best-fit model, or subset of models, was identified from a set of six candidate base models (Buckland et al. 2001) that included all pair-wise combinations of two key functions (half-normal or hazard-rate) and three series expansions (cosine, hermite polynomial, simple polynomial). In order to test whether these models could be improved by accounting for heterogeneity in detectability in relation to observer and environmental variables (Marques et al. 2007), the following covariates were then added to the best-fit base model(s) independently: wind speed, cue (whether the bird was seen or heard first), observer identity, management panel, survey number (where applicable) and, for sheep & beef and dairy sectors only, habitat type. Finally, for species where at least one covariate improved the model fit, pair-wise combinations of observer and environmental variables were tested to see if model fit could be improved further. For all three steps the best-fit model, or subset of candidate models, was identified using both the small sample size adaptation of Akaike’s Information Criterion (AICc) and visual inspection of the detection functions (Weller et al. 2012). Visual inspections of the best-fit models were carried out because detection functions were sometimes over-fitted, particularly when sample sizes were small. The latter were usually associated with an upward spike close to zero, resulting in overestimates of $f(0)$ and density. In these cases the next best fitting model without this issue was selected to estimate density. For all models, we reduced the risk of generating unrealistically large flock sizes and high variance density estimates by: (a) log-transforming flock sizes prior to fitting a regression model for observed flock size and distance from the transect; and (b) only estimating flock size using regression models that met a significance *P*-value threshold of 0.15, otherwise the mean was applied (Buckland et al. 2001).

Estimating bird densities

For each global and survey-specific dataset, an overall mean (weighted according to the sampling effort on each farm) and individual-farm mean density estimates for each survey, along with the corresponding coefficients of variance estimates, were extracted from the best-fitting detection function using Distance’s post-stratification features. Taking a model-averaging approach might result in systematic bias among species and sectors (Burnham & Anderson 1998). Therefore density estimates were extracted for a subset of species using both the best-fit and modelling averaging approaches and compared with those from the corresponding best-fit base model.

On sheep & beef and dairy farms, where it was difficult to identify the exact location of birds relative to the transect end-point, observers often recorded birds located beyond the end-point. It was not possible to exclude these observations from our analyses because observer location was not recorded. Thus, contrary to the distance-sampling assumption that the effective survey area boundary lies perpendicular with the end-point, it extended beyond the end-point in a semi-circle. To down-weight density estimates accordingly, a factor that took into account the species-survey-specific estimates of effective survey width and number of transects on each property was calculated: $(\Sigma_n (l_n \times esw_s \times 2) + \Sigma_n (pi \times esw_s^2)) / \Sigma_n (l_n \times esw_s \times 2)$, where n is number of transects on the property, l_n is the length of transect n, and esw_s is the effective survey width for species s. Density estimates for each property and survey were then divided by the corresponding correction factor.

Data analysis

Paired t-tests were used to compare median global and survey-specific estimates of effective survey widths as well as mean and CV density estimates among species within the kiwifruit and sheep & beef sectors. Linear mixed-effects models were used to test whether (1) effective survey widths (ESW) and (2) the precision (measured using coefficients of variance, CV) of density estimates varied predictably in relation to species' conspicuousness, number of detections recorded, and the complexity of the detection functions fitted (Bates & Maechler 2010). Only data from the survey-specific datasets were considered in these analyses. The best-fit model or subset of models from the four candidate models (null, conspicuousness or model-fit variables or both) were assessed using AICc values (quantified using the Laplace approximation). The 'conspicuousness' models included the following variables: body mass (as an index of body size; Heather & Robertson 2000), flocking behaviour (two-level categorical variable: solitary/pair vs group/flock/colony; Heather & Robertson 2000), sector (as a surrogate measure of habitat composition and structure) and conservation status (i.e. native or introduced). The 'model-fit' models included variables measuring the complexity of the detection function fitted (i.e. the total number of key function, adjustment terms and covariates included in the best-fit detection-function model) and the number of detections recorded. The number of detections and body mass were log-transformed prior to analysis. The response variables were the effective survey widths and the coefficient of variances for the overall mean density estimate for each species/sector/survey dataset (as extracted from Distance). To account for repeated measures of the same species across sectors and surveys, species identity was specified as a random effect. The relative importance of variables within the best-fit model was then assessed by a comparison of the parameter estimates. These analyses were carried out using the statistical package R, version 2.10.1 (R Development Core Team 2009).

Results

Survey effort

The number of sites in each survey was relatively stable in the kiwifruit sector over the 6-year study period, with only one site excluded from the third survey (Table 1). Over the same period, however, a large number of study sites were excluded from the sheep & beef sector, with the median number of study sites per cluster decreasing from three to two (Table 1). The loss of sheep & beef properties was reasonably evenly distributed among the different management panels and clusters, with only one complete cluster (in Marlborough) removed from the study altogether. Changes in ownership or conversion to other land use types or both were the main reasons why farms left the ARGOS study panel. Potential damage to arable crops by observers was another reason for farmers declining access to a small number of properties.

Overall, variation in survey effort among sectors was roughly proportional to the property sizes (Table 1). Thus, in the sheep & beef sector, the total length of transects per property and across all properties was on average 2–6 times longer than in the other sectors. Survey effort (number and length of transects) also varied over time, declining in kiwifruit and peaking in the second survey in the sheep & beef sector.

Number of species

Overall, detection functions were fitted for 33 (52%) of the 64 species detected across the three sectors, including 15 native species (Fig. 2). However, of those species, only nine introduced passerines and four native species (fantail, grey warbler, silvereye and welcome swallow) were common to all three sectors (Fig. 2).

The number of species recorded per survey was consistently highest in the sheep & beef sector, with the number of species detected per survey within each sector (for sequential surveys: $n_{DY} = 33$; $n_{KF} = 37, 38, 27$; $n_{SB} = 43, 49, 44$) increasing roughly in proportion to sampling effort (Table 1). Within each sector, ≤50% of the species observed per survey had ≤40 detections records, i.e. met the suggested sample-size threshold for fitting detection functions (percentage of species for sequential surveys: DY = 50; KF = 46, 34, 26; SB= 53, 47, 41). Thus, the number of species for which detection functions were fitted (using the survey-specific datasets) varied among sectors (total: $n_{SB} = 23$, $n_{KF} = 16$; $n_{DY} = 17$) and surveys (range: $n_{SB} = 19$–23, $n_{KF} = 11$–16), with the numbers also increasing roughly in proportion to survey effort (Table 1; Appendices 1–3).

Using the global dataset (i.e. all three surveys combined), the total number of species for which detection functions could be fitted increased (relative to the survey-specific datasets) by 17% and 31% in sheep & beef and kiwifruit sectors respectively ($n_{SB} = 27$, $n_{KF} = 21$; Figs 2–4). When the number of sheep & beef sites was reduced (based on the 26SB dataset), the number of species (with detection functions fitted) also increased by 9% for the global dataset (24 species) relative to the survey-specific ones (22 species; Appendix 4).

For 11 species in the kiwifruit sector, detection functions could be fitted for all three surveys but only half of those (all introduced passerines) consistently had datasets with >80 detections (Figs 2–4; Appendix 2). For the other five species (mainly natives), the number of detections recorded in the last survey (when survey effort per property was lower; Table 1) was below the recommend threshold (≤40 detections; Appendix 2). For fantail and kingfisher, there were only a few detection records (25–26) collected in the third survey, so only base detection functions (i.e. without any covariates) could be fitted. In the kiwifruit global dataset, detection functions were fitted for species with ≥ 50 detections (Appendix 2).

In the sheep & beef sector, three survey-specific detection functions could be modelled for 19 species (including seven native ones) irrespective of the number of study sites considered (Figs 2–4). However, detection functions for harrier and silvereye in the last survey (when there were only 26 study sites) were based on small datasets (≤40 detections; Appendix 3). Two survey-specific detection functions were also derived for the other four species in the full dataset (Figs 2–4), albeit

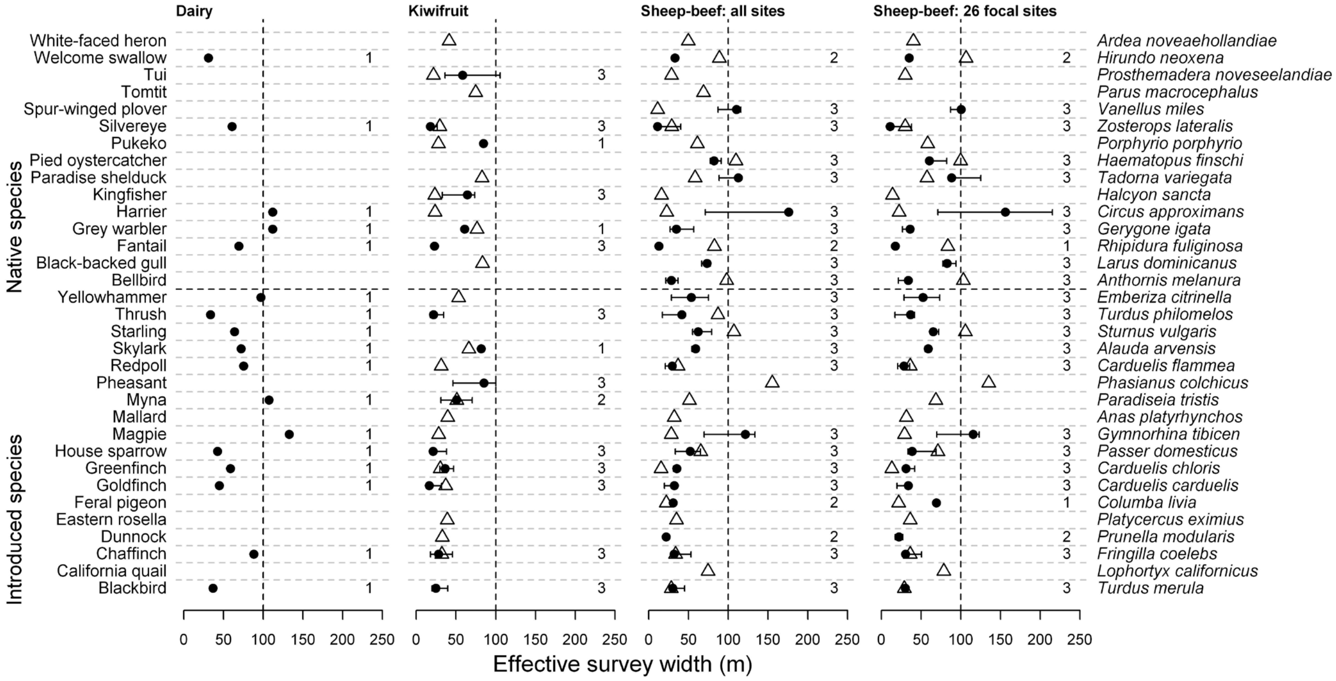

Figure 2. Summary of effective survey widths in relation to bird species, agricultural sector, and survey in New Zealand. Each plot shows the median effective survey width (and minimum–maximum range) from survey-specific detection functions (circles with bars) as well as the estimates from the global detection functions (triangles). Numbers on the right-hand side of each plot indicate the number of survey-specific detection functions calculated per species (see Appendices 1–4 for full list of models). (Note: in the kiwifruit sector, bird locations were recorded in bands up to 100 m from the transect line.)

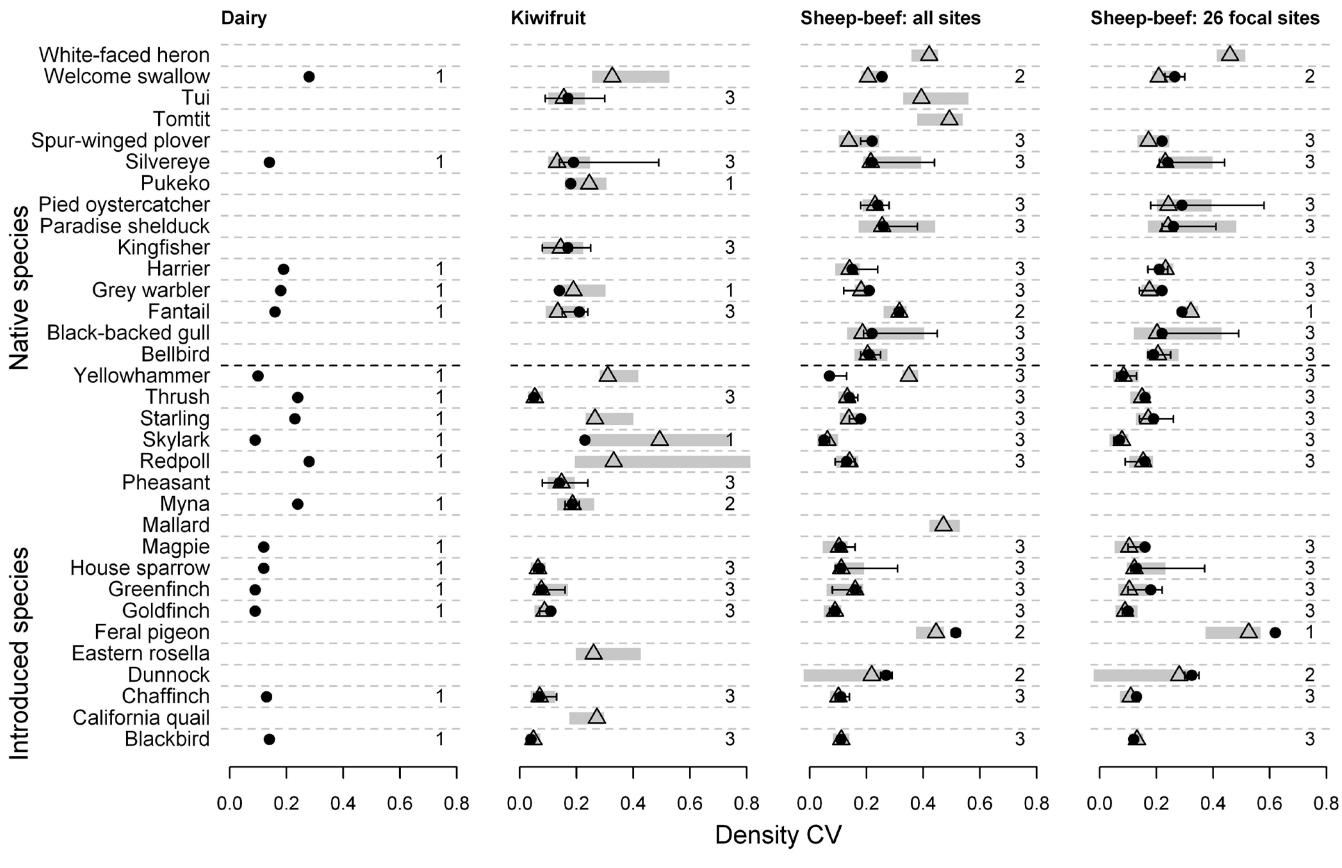

Figure 3. Summary of survey-level density CV estimates in relation to bird species and agricultural sector in New Zealand. Each plot shows the median density CV (and minimum–maximum range) from survey-specific (circles with thin black bars) and global detection functions (triangles with thick grey bars). Numbers on the right-hand side of each plot indicate the number of survey-specific detection functions calculated per species (see Appendices 1–4 for full list of models).

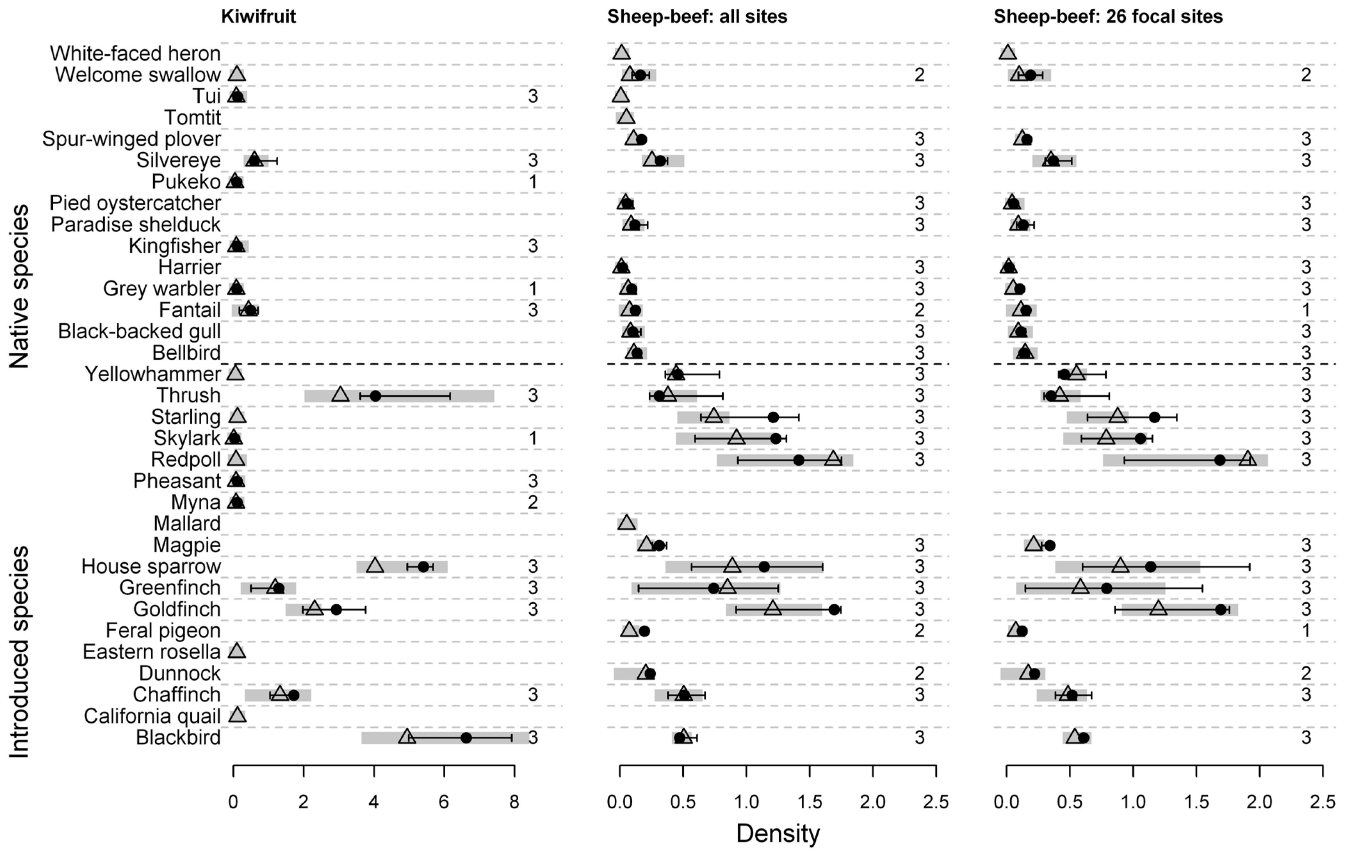

Figure 4. Comparisons of survey-level density estimates, calculated using the global and survey-specific detection functions for bird species within New Zealand kiwifruit and sheep & beef sectors. Each plot shows the median density (and minimum–maximum range) from survey-specific (circles with thin black bars) and global detection functions (triangles with thick grey bars). Numbers on the right-hand side of each plot indicate the number of survey-specific detection functions calculated per species (see Appendices 1–4 for full list of models).

from datasets with small sample sizes for feral pigeon and fantail in at least one survey (Appendix 3). The latter two species were either excluded or had only one survey-level estimate for the 26SB dataset (Figs 2–4; Appendix 4). Detection functions fitted to species in the global dataset were all based on relatively large sample sizes (69–3774 detections), except for tomtit, tūī and white-faced heron ($n \leq 52$; Appendix 3).

Detection functions

Global estimates of effective survey widths were comparable with matched median estimates from the survey-specific datasets for the same species within the sheep & beef sector (Fig. 2; *t*-test, *P*-values > 0.06), but significantly higher in the kiwifruit sector ($t = -2.3$, $P = 0.037$). However, while the probability of detecting some species was similar among the different sectors and surveys (e.g. blackbird and skylark), it was highly variable for others either over time (e.g. harrier, tūī and pheasant) or among sectors (e.g. yellowhammer). Both conspicuousness and model-fit variables were included in the best-fit model accounting for variation in effective survey widths (Appendix 5; Fig. 5), with wider effective survey widths associated with: (1) larger bodied species; (2) native species (relative to introduced ones); (3) the dairy sector (relative to the sheep & beef and kiwifruit sectors); and (4) higher numbers of detection records. Effective survey widths for the same species in the different sectors were only comparable for a few species (i.e. blackbird, chaffinch and greenfinch in the sheep & beef and kiwifruit sectors; magpie, starling and welcome swallow in the sheep & beef and dairy sectors; Appendix 6).

Covariates in best-fit detection functions

For the survey-specific datasets, at least 58% of the best-fit detection functions in each sector included at least one covariate (min–max range: DY = 88%; KF = 58–75%; SB = 58–74%). For all three sectors, the most common covariates were the cue (seen or heard) used by observers to detect birds (DY = 53%; KF = 43–58%; SB = 32–61%) and observer identity (DY = 18%; KF = 17–47%; SB =13–32%). Wind and panel were retained in a smaller proportion of models (Wind: DY = 6%; KF = 8–20%; SB = 5–30%; Panel: DY = 18%; KF = 0–9%; SB = 5–9%). Habitat was only considered as a potential covariate for sheep & beef and dairy farms, where it was retained in 4–9% and 6% of the best-fit models respectively.

Comparing base, best-fit and model-averaged detection functions

Density and ESW estimates (and CVs) from the best-fit model were comparable, if not identical, to those extracted using a model-averaging approach (based on the subset of best-fit models) for six focal species (blackbird, magpie, skylark,

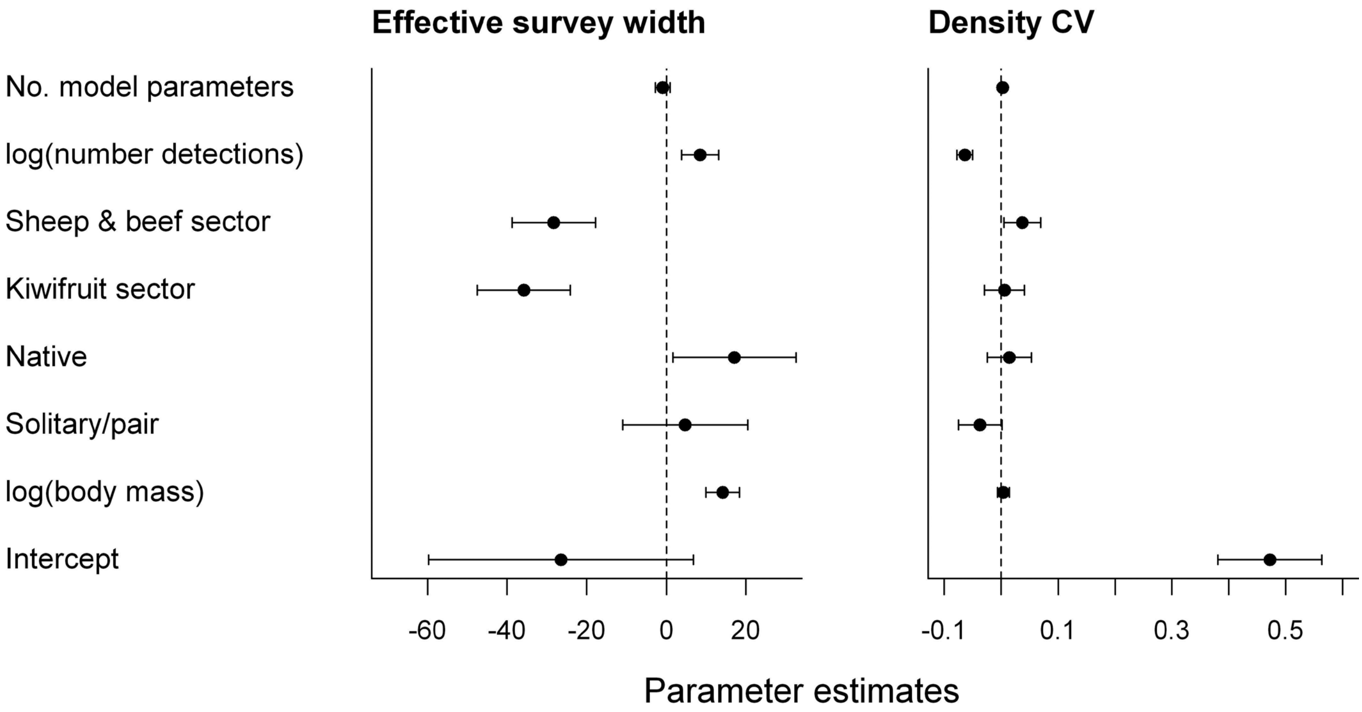

Figure 5. Parameter estimates for the 'best-fit' (linear mixed-effects regression) model (Appendix 5) for birds in New Zealand agricultural sectors, explaining variation in effective survey widths as well as coefficient of variance (CV) estimate for densities extracted from the survey-specific detection functions.

goldfinch, grey warbler and fantail) in two sectors (sheep & beef and kiwifruit), irrespective of the datasets considered (global or survey level; Appendix 7). However, density estimates extracted from the base models were usually slightly lower than the corresponding best-fit and model-averaged estimates, while the reverse pattern was observed for the effective survey widths. Overall, coefficient of variance estimates for densities were similar for most species. In addition, the overall best-fit model captured the predominant covariates identified in the subset of best-fit models used for model-averaging (Appendix 7).

Precision of density estimates

For all datasets, except SB26, coefficient of variances for survey-level density estimates extracted from global detection functions were comparable with matched median estimates from survey-specific models (Fig. 3; *t*-test; d.f. = 10–22, *P*-values > 0.05). For two out of the three surveys in the focal SB26 dataset, CV estimates extracted from the survey-specific detection functions were significantly higher than those from global models (*t*-test; d.f. = 18–21 ; *P*-values < 0.02). Reducing the spatial resolution of the sheep & beef dataset also increased the density CV estimates for the survey-specific models (*t*-test, d.f. = 21, *P*-values < 0.005) but not the global ones (*t*-test; d.f. = 23, *P*-values > 0.05).

Variation in density CV estimates extracted from the survey-specific detection functions was related to variables associated with both conspicuousness and model-fit (Appendix 5). More precise measures of density (i.e. low CV estimates) were associated with higher numbers of detections, while less precise ones were more likely for the sheep & beef sector and species exhibiting flocking or colonial nesting behaviour (e.g. black-backed gulls and feral pigeons; Fig. 5; Appendix 4).

Comparing density estimates

Overall, survey-level density estimates extracted from the global detection functions were highly correlated with corresponding estimates from the survey-specific detection functions in the kiwifruit and sheep & beef sectors (Spearman's correlation coefficient range: ρ_S = 0.97–0.99, *P* < 0.001; Fig. 4). However, density estimates calculated using the global detection function for species within the kiwifruit sector were significantly lower than matched estimates from the survey-specific detection functions in the initial survey (*t* = −2.38, d.f. = 14, *P*-value = 0.03) but higher in the two subsequent surveys (*t* = 2.61–2.96, d.f. = 10–15, *P*-value < 0.03). In the sheep & beef sector, irrespective of the dataset considered (SB or 26SB), density estimates based on the global detection function were comparable with those from the survey-specific functions for the initial two surveys (*t* = 1.05–1.82, d.f. = 21–22, *P* > 0.05; Fig. 4), but not the third survey, when they were significantly lower (*t* = 4.54–4.47, d.f. = 18, *P* < 0.001) Also, altering the spatial resolution of the detection function for the sheep & beef data (i.e. SB vs 26SB) did not affect median species density estimates for the 26 focal farms (when comparing matching global and survey-specific datasets: *t*-test, *P* > 0.05) for all surveys, except the second survey where the smaller dataset overestimated densities relative to the larger dataset when using the global detection functions to estimate density.

Discussion

Design considerations for long-term monitoring schemes

To determine whether the level of sampling effort proposed for a monitoring scheme design is appropriate to meet its objectives, a pilot study can be useful. For example, based on data collected from three surveys, the ARGOS scheme was only able to extract density estimates for ≤ 50% of species recorded and these were primarily introduced species. If quantifying changes in the density of native species over time is identified as a research priority, then a more intensive sampling effort per study site or an increase in the number of study sites would be required to calculate robust measures of density.

To overcome the limitation of small sample sizes for some species in our analysis, we pooled data collected over a number of years within each sector to fit a global detection function. If exactly the same bird survey protocol had been implemented in all three sectors and surveys, then these data could have been pooled to fit a global detection function (with each survey–sector combination specified as a strata) to measure and account for detection probabilities of a wider range of species. Alternatively, other metrics (e.g. site occupancy) and analysis tools could be employed to monitor the status of native species' populations (e.g. MacKenzie 2005; MacKenzie & Royle 2005; Alldredge et al. 2007a).

Strategies for dealing with loss or addition of study sites or changes in sampling protocols in long-term monitoring schemes are also important design considerations. Such strategies need to take into account the scheme's objectives to mitigate the risk of introducing sampling bias and error associated with any modifications to the scheme (e.g. Newson et al. 2005, 2008). In the case of the ARGOS bird monitoring scheme, replacing lost sheep & beef study sites with ones from similar localities and with similar management regimes would facilitate development of a longitudinal dataset for monitoring the impact of land management changes within that sector. But expanding the spatial zone of inference of the dataset by introducing new study sites from outside the current study region (e.g. on the North Island) or with different management strategies will not add to the power of the longitudinal dataset because the spatial and temporal variables would be confounded (e.g. with the introduction of new species and climatic variables). For example, exclusion from some arable sites in the later ARGOS sheep & beef farm surveys means that sampling effort was probably biased towards pastoral-dominated sites. Such sampling bias needs due consideration when interpreting any temporal changes in bird abundance among surveys, by either down-weighting observations from arable-dominated sites in the earlier surveys or excluding those arable sites from any longitudinal comparisons and acknowledging the reduction in the zone of inference of the study.

Where sustaining a constant level of sampling effort over time is not feasible, any variation in sampling effort at study sites over time needs to be taken into account when analysing and interpreting the data. In the ARGOS surveys, both the number of species detected and the number of detections per species were roughly correlated with survey effort. In the third survey of the kiwifruit sector, for example, only one circuit of each study site was completed, thus reducing the number of species and the number of detections per species recorded relative to earlier more intensive surveys. At least two circuits per site per survey are needed to obtain sufficient detection records for native species across the 36 kiwifruit study sites.

Loss of one-third of the ARGOS sheep & beef study sites is also likely to reduce the scheme's power to detect changes in bird community composition and abundance in relation to changes in land management systems in this sector. The ARGOS surveys were carried out over a long time-frame each day (0800–1600 hours) to increase the time available to conduct surveys, so future analyses should test for variation in bird activity associated with the timing of surveys (Weller et al. 2012).

Importance of explicit protocols for field survey, data management and analysis

Cost-effective designs for bird monitoring schemes need to optimise both sampling effort in the field and time spent managing and analysing data. Rigorous training programmes are required, to ensure not only that field teams record high quality data in the field but also that those data are then checked, entered and edited following explicit protocols (e.g. guidelines for removing from the dataset or editing records with missing information). Similar processes should be in place for data analysis (e.g. protocols for fitting detection functions when extracting density estimates) to minimise the risk of bias. Clear documentation of any changes to such protocols reduces the risk of loss of 'institutional knowledge' associated with changes in field teams or staff analysing and interpreting the data. For example, contrary to the distance-sampling assumption that the effective survey area boundary lies perpendicular with the end-point, ARGOS observers often recorded birds beyond the end-point in a semi-circle for the sheep & beef and dairy sectors. While field protocols can be adjusted to measure this error (e.g. by using GPS to record the observer's location and using that information to plot the location of birds), altering the field protocols in this way means it is no longer appropriate to pool data across all surveys to fit global detection functions.

Pros and cons of different analytical approaches

To facilitate a more cost effective process for data analysis, density estimates for the ARGOS bird monitoring scheme were extracted using the best-fit model rather than the model-averaging approach. This approach seems valid as, for the subset of species considered, no systematic differences were observed in the detection function covariates or estimates of ESW and density extracted from the two models.

Altering the temporal resolution of the datasets used to fit detection functions did influence the number and precision of density estimates that could be extracted. As predicted, within each sector, pooling data from all surveys increased the number and precision of density estimates that could be extracted, by increasing the number of detections per species. However, while density estimates for matched global and survey-specific datasets were highly correlated in the sheep & beef and kiwifruit sectors, they were not comparable for all surveys. Also, within the sheep & beef sector, varying the spatial resolution of the datasets used to fit detection functions did alter the density estimates or the precision of those estimates for the subset of 26 focal study sites. Any comparisons of density estimates extracted from global and survey-specific datasets need to take these factors into consideration.

Importance of accounting for heterogeneity in detectability

Distance sampling, which was implemented as part of the ARGOS bird monitoring scheme, is one technique used to estimate the probability of detection for individual species and produce more accurate measures of density. However, Alldredge et al. (2007b, 2008) emphasised the practical difficulties of implementing distance-sampling techniques accurately in the field. Using a song playback system in a forest habitat, Alldredge et al. (2007b) found a non-linear relationship between measurement error and distance; although measure error was substantial, it could be reduced by 15% with training. The observed changes in effective survey widths for the same species over space and time in the ARGOS scheme (Figs 2–5) could be driven by either changes in observer or bird behaviour.

Such variation in detectability is often overlooked when making comparisons of raw counts among species, sectors and surveys (Buckland et al. 2001; Norvell et al. 2003). For example, the five-minute bird count method (5MBC; Dawson & Bull 1975), the most commonly used bird monitoring technique in New Zealand (Hartley 2012), assumes detection probabilities are constant among species and habitats. Our analyses indicate that this assumption is likely to result in biased estimates of abundance and species diversity (as diversity measures also assume that abundance estimates for different species and habitats are comparable), with observer variation added and subsumed within the CV estimates. If the main agenda is to detect long-term trends in abundance in the same places and habitats that do not change much (or at least in ways that affect detectability), bird counts can be treated as relative indices of abundance and reliable indicators of trends. However, if the habitat mosaics in production landscapes change, as urged for ecological restoration itself, then relative indices of abundance may mislead farmers and restoration managers because detection probability, as well as putative abundance and diversity changes, will affect monitoring results.

The importance of accounting for heterogeneity in detectability in relation to observer and environmental variables was also highlighted in this paper, with most best-fit detection functions including at least one covariate. For example, the cue used to detect birds was the most common covariate in the ARGOS scheme, showing that raw counts among species, habitats and time periods that assume equal detection probabilities for seen and heard observations (e.g. 5MBCs) are likely to be biased. Conversely, management panel and habitat composition were rarely selected as covariates when fitting detection functions. This suggests that if these variables within each sector are not correlated then management panel is a relatively unimportant feature influencing detection probabilities (supporting the findings of Weller et al. (2012)). Provided that all other environmental and observer variables remained unchanged, raw counts could then be used as an index for monitoring bird populations on the ARGOS study sites.

The importance of training observers for bird monitoring schemes, which other studies have previously demonstrated (e.g. Alldredge et al. 2007b), was reiterated in the ARGOS scheme, with observer identity included in a high proportion of the best-fit detection functions. In the ARGOS bird monitoring scheme, budget constraints meant that the field teams for each survey consisted primarily of novice observers. While these observers were subject to an intensive training programme at the start of each field season, it is likely that cryptic species were either misidentified or missed altogether particularly early in the season. Weller et al. (2012), for example, found that blackbird encounter rates on sheep & beef farms were similar for both experienced and inexperienced observers, but song thrush encounter rates were lower for inexperienced observers relative to experienced ones. Future studies investigating longitudinal changes in species composition on the ARGOS study sites need to take such limitations into consideration. The risk of observer bias in long-term monitoring schemes can be mitigated, to some extent, by either having a small group of well-trained observers or a bigger sample of potential observers who are randomly distributed among the sites. Otherwise, observer identity should be accounted for during analyses of any raw count data. The contribution of 'citizen science', where large numbers of amateur enthusiasts are enlisted to help monitor bird populations, is immense overseas (e.g. Baillie 1990; Gregory et al. 2004, 2005). OSNZ's bird mapping scheme (Robertson et al. 2007) is a notable contribution already, and the recently instigated garden bird surveys (Spurr 2012) will soon contribute nationally.

Conclusions

In the absence of information on bird abundance and variability and sampling precision, the overall ARGOS sampling effort was determined mainly by practical and financial resourcing constraints. Power analyses are now needed to assess the feasibility of detecting different levels of change in long-term population trends (Eaton et al. 2009) at the sector level, and especially for all of New Zealand's agricultural landscapes, using the current and alternative survey designs. Despite their analytical and practical challenges, use of distance methods will help account for the differential detection probabilities already demonstrated. In the meantime less than half the birds present can be modelled using distance methods, but gradually more species can be added as sufficient detections are gathered to generate their global detection functions. This accumulation of data will also allow gradual reduction in sampling uncertainty for key species and likely increased power to detect any ongoing declines. The national need for monitoring birds in production areas is unlikely to be met by a single and normal research team budget, so more national coordination and standardisation of methods based on our experience is now required.

Acknowledgements

This work was funded by the Foundation for Research, Science and Technology (Contract No. AGRB0301) with additional assistance from the Certified Organic Kiwifruit Producers Association, Fonterra, Merino New Zealand Inc., a meat packing company, ZESPRI Innovation Company, the University of Otago, and Landcare Research. For assistance with the fieldwork, we thank Jayson Benge, Dave Lucock, Corey Bragg, Dean Clarke, Erin O'Neill, Francesca Buzzi, Guinny Coleman, Jennifer Techer, Jeroen Lurling, Jess Seipe, Jessica Lach, Jo Wright, Libby Mckinnel, Logan Penniket, Natasha Percey, Steve Rate, Susan Mcguire and Tracey Dearlove. We are grateful to Richard Gregory and David Baker for their helpful reviews and Christine Bezar and Dan Tompkins for their editorial input.

References

Alldredge MW, Pollock KH, Simons TR, Shriner SA 2007a. Multiple-species analysis of point count data: a more parsimonious modelling framework. Journal of Applied Ecology 44: 281–290.

Alldredge MW, Simons TR, Pollock KH 2007b. A field evaluation of distance measurement error in auditory point count surveys. Journal of Wildlife Management 71: 2759–2766.

Alldredge MW, Pacifici K, Simons TR, Pollock KH 2008. A novel field evaluation of the effectiveness of distance and independent observer sampling to estimate aural avian detection probabilities. Journal of Applied Ecology 45: 1349–1356.

Baillie SR 1990. Integrated population monitoring of breeding birds in Britain and Ireland. Ibis 132: 151–166.

Bates D, Maechler M 2010. lme4: Linear mixed-effects models using S4 classes. R package version 0.999375-33. http://CRAN.R-project.org/package=lme4

Benton TG, Bryant DM, Cole L, Crick HQP 2002. Linking agricultural practice to insect and bird populations: a historical study over three decades. Journal of Applied Ecology 39: 673–687.

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L 2001. Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press. 432 p.

Burnham KP, Anderson DR 1998. Model selection and inference: a practical information-theoretic approach. New York, Springer. 353 p.

Butler SJ, Vickery JA, Norris K 2007. Farmland biodiversity and the footprint of agriculture. Science 315: 381–384.

Carey PL, Benge JR, Haynes RJ 2009. Comparison of soil quality and nutrient budgets between organic and conventional kiwifruit orchards. Agriculture, Ecosystems & Environment 132: 7–15.

Chamberlain DE, Fuller RJ, Bunce RGH, Duckworth JC, Shrubb M 2000. Changes in the abundance of farmland birds in relation to the timing of agricultural intensification in England and Wales. Journal of Applied Ecology 37: 771–788.

Dawson DG, Bull PC 1975. Counting birds in New Zealand forests. Notornis 22: 101–109.

DEFRA 2002. Achieving a better quality of life: a review of progress towards sustainable development. London, Department for the Environment, Food & Rural Affairs.

Donald PF, Green RE, Heath MF 2001. Agricultural intensification and the collapse of Europe's farmland bird populations. Proceedings of the Royal Society of London series B 268: 25–29.

Eaton MA, Brown AF, Noble DG, Musgrove AJ, Hearn R, Aebischer NJ, Gibbons DW, Evans A, Gregory RD 2009. Birds of Conservation Concern 3: the population status of birds in the United Kingdom, Channel Islands and the Isle of Man. British Birds 102: 296–341.

Foley JA, DeFries R, Asner GP, Barford C, Bonan G, Carpenter SR, Chapin FS, Coe MT, Daily GC, Gibbs HK, Helkowski JH, Holloway T, Howard EA, Kucharik CJ, Monfreda C, Patz JA, Prentice IC, Ramankutty N, Snyder PK 2005. Global consequences of land use. Science 309: 570–574.

Fewster RM, Buckland ST; Siriwardena GM, Baillie SR, Wilson JD 2000. Analysis of population trends for farmland birds using generalized additive models. Ecology 81: 1970–1984.

Fuller, RJ, Gregory, RD, Gibbons DW, Marchant JH, Wilson JD, Baillie SR, Carter N 1995. Population declines and range contractions among lowland farmland birds in Britain. Conservation Biology 9: 1425–1441.

Furness RW, Greenwood JJD (eds) 1993. Birds as monitors of environmental change. London, Chapman & Hall.

Gibbons DW, Reid JB, Chapman RA (comps) 1993. The new atlas of breeding birds in Britain and Ireland, 1988–1991. London, T & AD Poyser. 520 p.

Golden JS, Dooley KJ, Anderies JM, Thompson BH, Gereffi G, Pratson L 2010. Sustainable product indexing: navigating the challenge of ecolabeling. Ecology and Society 15(3): 8 [online]. URL: http://www.ecologyandsociety.org/vol15/iss3/art8.

Gregory RD, Noble DG, Custance J 2004. The state of play of farmland birds: population trends and conservation status of lowland farmland birds in the United Kingdom. Ibis 146 (Suppl. s2): 1–13.

Gregory RD, van Strien A, Vorisek P, Gmelig Meyling AW, Noble DG, Foppen RPB, Gibbons DW 2005. Developing indicators for European birds. Philosophical Transactions of the Royal Society Series B 360: 269–288.

Hartley LJ 2012. Five-minute bird counts in New Zealand. New Zealand Journal of Ecology 36: 268-278.

Heather BD, Robertson HA 2000. The field guide to the birds of New Zealand. Rev. edn. Auckland, Viking. 440 p.

Hole DG, Whittingham MJ, Bradbury RB, Anderson GQA, Lee PLM, Wilson JD, Krebs JR 2002. Agriculture: Widespread local house-sparrow extinctions. Nature 418: 931–932.

Hole DG, Perkins AJ, Wilson JD, Alexander IH, Grice PV, Evans AD 2005. Does organic farming benefit biodiversity? Biological Conservation 122: 113–130.

Krebs JR, Wilson JD, Bradbury RB, Siriwardena GM 1999. The second Silent Spring? Nature 400: 611–612.

MacKenzie DI 2005. What are the issues with presence–absence data for wildlife managers? Journal of Wildlife Management 69: 849–860.

MacKenzie DI, Royle JA 2005. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42: 1105–1114.

MacLeod CJ, Moller H 2006. Intensification and diversification of New Zealand agriculture since 1960: an evaluation of current indicators of land use change. Agriculture, Ecosystems & Environment 115: 201–218.

MacLeod CJ, Blackwell G, Moller H, Innes J, Powlesland R 2008. The forgotten 60%: bird ecology and management in New Zealand's agricultural landscape. New Zealand Journal of Ecology 32: 240–255.

MacLeod CJ, Newson SE, Blackwell G, Duncan RP 2009. Enhanced niche opportunities: can they explain the success of New Zealand's introduced bird species? Diversity & Distributions 15: 41–49.

MacLeod CJ, Tompkins, DM, Drew KW, Pyke N 2011. Does farm-level habitat composition predict pest-bird numbers and distribution? Wildlife Research 38: 464–474.

MacLeod CJ, Blackwell G, Benge J 2012. Reduced pesticide toxicity and increased woody vegetation cover account for enhanced native bird densities on organic orchards. Journal of Applied Ecology 49: 652–660.

Marques TA, Thomas L, Fancy SG, Buckland ST 2007. Improving estimates of bird density using multiple-covariate distance sampling. The Auk 124: 1229–1243.

Meurk CD, Swaffield SR 2000. A landscape ecological framework for indigenous regeneration in rural New Zealand-Aotearoa. Landscape and Urban Planning 50: 129–144.

Moller H, Wearing A, Pearson A, Perley C, Steven D, Blackwell G, Reid J, Johnson M 2005. Environmental monitoring and research for improved resilience on ARGOS farms. ARGOS Working Paper No. 6. Dunedin, University of Otago. 136 p. http//www.argos.org.nz

Moller H, MacLeod CJ, Haggerty J, Rosin C, Blackwell G, Perley C, Meadows S, Weller F, Gradwohl M 2008.

Intensification of New Zealand agriculture: implications for biodiversity. New Zealand Journal of Agricultural Research 51: 253–263.

Newson SE, Woodburn RJW, Noble DG, Baillie SR, Gregory RD 2005. Evaluating the Breeding Bird Survey for producing national population size and density estimates. Bird Study 52: 42–54.

Newson SE, Evans KL, Noble DG, Greenwood JJD, Gaston KJ 2008. Use of distance sampling to improve estimates of national population sizes for common and widespread breeding birds in the UK. Journal of Applied Ecology 45: 1330–1338.

Norton DA, Miller CJ 2000. Some issues and options for the conservation of native biodiversity in rural New Zealand. Ecological Management & Restoration 1: 26–34.

Norvell RE, Howe FP, Parrish JR 2003. A seven-year comparison of relative-abundance and distance-sampling methods. The Auk 120: 1013–1028.

PCE (Parliamentary Commissioner for the Environment) 2004. Growing for good: Intensive farming, sustainability and New Zealand's environment. Wellington, Parliamentary Commissioner for the Environment. 236 p.

PCE (Parliamentary Commissioner for the Environment) 2010. How clean is New Zealand? Measuring and reporting on the health of our environment. Wellington, Parliamentary Commissioner for the Environment.

Peach WJ, Lovett LJ, Wotton SR, Jeffs C 2001. Countryside stewardship delivers cirl buntings (*Emberiza cirlus*) in Devon, UK. Biological Conservation 101: 361–373.

Perley C, Moller H, Hamilton WJ, Hutcheson J 2001. Towards safeguarding New Zealand's agricultural biodiversity: research gaps, priorities and potential case studies. Ecosystem Consultants Report 23: 1–230. http://ww.maf.govt.nz/mafnet/rural-nz/sustainable-resource-use/biodiversity/convention-ion-biodiversity/cbd-report.pdf

R Development Core Team 2009. R: A language and environment for statistical computing. Vienna, Austria, R Foundation for Statistical Computing. ISBN 3-900051-07-0, URL http://www.R-project.org.

Robertson CJR, Hyvönen P, Fraser MJ, Pickard CR 2007. Atlas of bird distribution in New Zealand 1999–2004. Wellington, The Ornithological Society of New Zealand. 533 p.

Siriwardena GM, Baillie SR, Buckland ST, Fewster RM, Marchant JH, Wilson JD 1998. Trends in the abundance of farmland birds: a quantitative comparison of smoothed Common Birds Census indices. Journal of Applied Ecology 35: 24–43.

Siriwardena GM, Baillie SR, Crick HQP, Wilson JD, Gates S 2000. The demography of lowland farmland birds. In: Aebischer NJ, Evans AD, Grice PV, Vickery JA eds Ecology and conservation of lowland farmland birds. Tring, Herts., UK, British Ornithologists' Union. Pp. 117–133.

Siriwardena GM, Stevens DK, Anderson GQA, Vickery JA, Calbrade NA, Dodd S 2007. The effect of supplementary winter seed food on breeding populations of farmland birds: evidence from two large-scale experiments. Journal of Applied Ecology 44: 920–932.

Sotherton NW, Self MJ 2000. Changes in plant and arthropod biodiversity on lowland farmland: an overview. Proceedings of the 1999 British Ornithologists' Union Spring Conference, University of Southampton, UK, 27–28 March 1999. Tring, Herts., UK, British Ornithologists' Union.

Spurr E 2012. New Zealand Garden Bird Survey – analysis of the first four years. New Zealand Journal of Ecology 36: 287-299.

Thomas L, Buckland ST, Rexstand EA, Laake JL, Strindberg S, Hedley SL, Bishop JRB, Marques TA, Burnham KP 2010. Distance software: design and analysis of distance sampling surveys for estimating population size. Journal of Applied Ecology 47: 5–14.

Tilman D 1999. Global environmental impacts of agricultural expansion: the need for sustainable and efficient practices. Proceedings of the National Academy of Sciences (USA) 96: 5995–6000.

Weller F 2012. A comparison of different approaches to monitoring bird density on New Zealand sheep & beef farms. New Zealand Journal of Ecology 36: 382-390.

Weller F, Blackwell G, Moller H 2012. Detection probability for estimating bird density on New Zealand sheep & beef farms. New Zealand Journal of Ecology 36: 371-381.