2025

Te Kura Pūtaiao Koiora

School of Biological Sciences

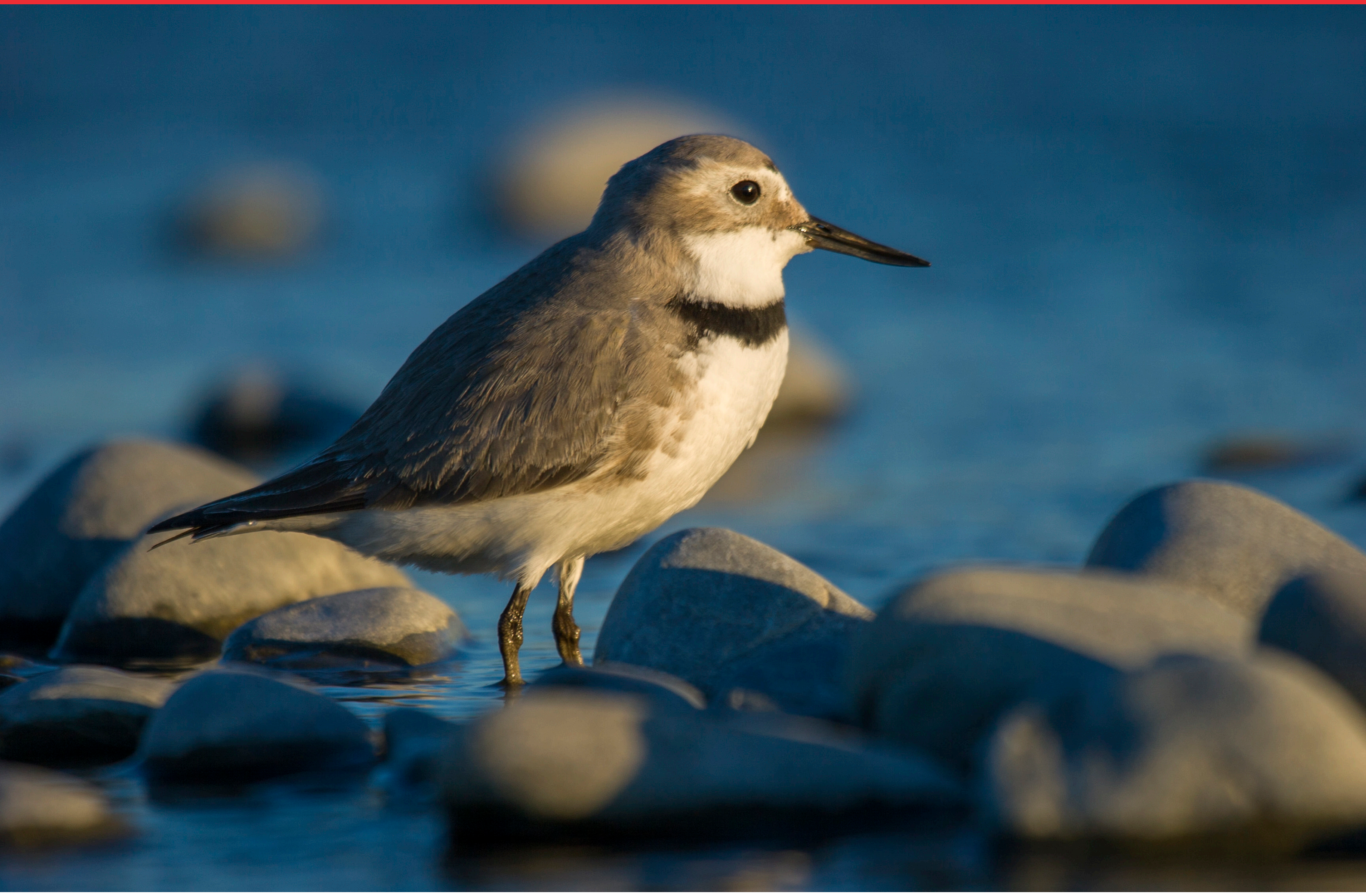
UNIVERSITY OF CANTERBURY

Te Whare Wānanga o Waitaha

CHRISTCHURCH NEW ZEALAND

Research report

Analysis of Upper Rakitata annual braided river bird walk-through survey data



UNIVERSITY OF CANTERBURY | Science Pūtaiao

Prepared for the Department of Conservation

Prepared by

Helen Warburton, Jack Anderson, Angus McIntosh

For further information about this report please contact the contributing authors

Te Whare Wānanga o Waitaha | University of Canterbury,
Private Bag 4800, Christchurch 8140

Email: helen.warburton@canterbury.ac.nz, angus.mcintosh@canterbury.ac.nz

Suggested citation

Warburton H. J., McIntosh A. R., Anderson, J. (2025). Analysis of Upper Rakitata braided river bird count data. Report prepared for Department of Conservation.
doi:10.6084/m9.figshare.30203944

Cover photograph

Wrybill | Ngutu pare nest on South Island braided riverbeds and the upper Rakitata is a national stronghold. Photo: Angus McIntosh

Copyright information

Disclaimer

Contents

Executive summary

Annual walk-through surveys on the Upper Rakitata River provide important information on threatened braided river birds such as wrybills and black-fronted terns. We analysed recent data (2021–2024) to check whether population trends could be detected and whether the survey approach is fit for long-term monitoring use.

Key findings

- Three species showed significant trends over time: southern black-backed gulls declined sharply (likely reflecting management actions to control their populations), wrybills declined more moderately raising concern, and black-billed gulls increased, though the latter was possibly influenced by an outlier in one section in 2023.
- For most species, large population changes (±10% per year) could potentially be detected with 10–15 years of surveys, but detecting smaller changes (±5% per year) will usually require 15–20 years.
- Statistical models of bird population trends that do not account for observer effort suggest stronger power to detect effects. However, these models risk confounding varying survey effort with real population change. Adjusting for effort is more conservative and likely more reliable for trend assessment.

Implications

The current annual walk-through surveys dataset provides a valuable foundation, but long-term, consistent monitoring will be needed to detect more subtle trends or predator control impacts. Improvements in survey design, consistent recording of effort and metadata, and consideration of complementary methods will strengthen the programme's ability to inform management decisions. Furthermore, attributing causes to population changes may be difficult given that the data are from a single catchment and population changes will take 10+ years to detect, so a combination of approaches will expand the strength of inference and value of the Upper Rakitata River bird surveys.

Background

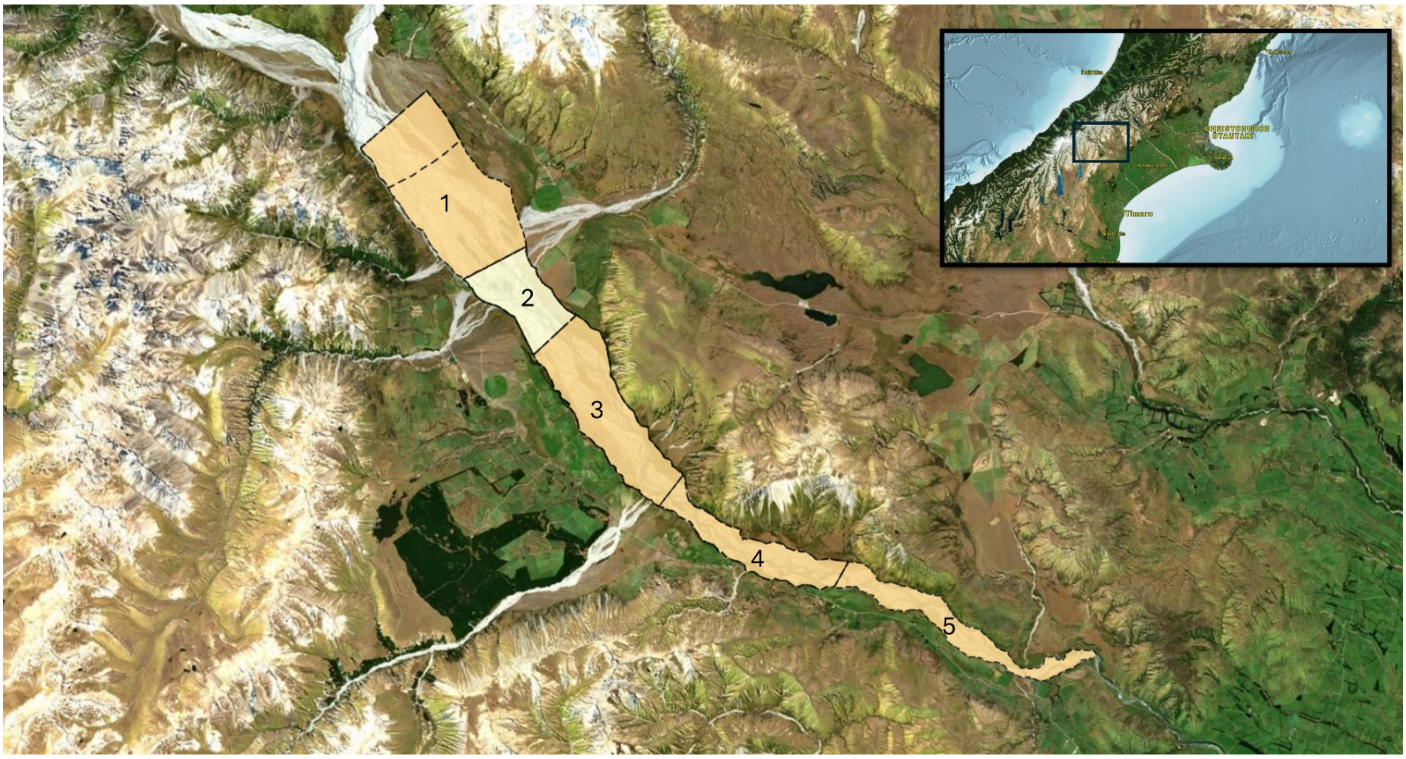


Figure 1 Map of the Upper Rakitata, from the confluence of the Clyde and Havelock rivers (top) to the Rakitata Gorge (bottom). Polygons of the current walk-through sections, 1 through to 5, are overlayed.

Sourced from https://ecan.maps.arcgis.com/home/item.html?id=c0247e7c6afd46c484b22d468b7ed39a

The Upper Rakitata River is valuable habitat to many threatened or at risk braided river bird species. Ngutu pare / wrybills (*Anarhynchus frontalis*) and tarapirohe / black-fronted terns (*Chlidonias albostriatus*) are two of the focal species driving the Department of Conservation's (DOC) Upper Rakitata Trapping project and are of particular interest for the annual Upper Rakitata walk-through survey of braided river birds (Figure 1). This walk-through had its inception in 1986 and has since been repeated various times, becoming an annual survey in 2020 (Figure 2). Given this awa provides significant breeding habitat for key threatened bird species (O'Donnell *et al.* 2016), it is important to determine whether DOCs efforts to support the populations of these braided river specialists are effective. This is the primary reason for the annual walk-through surveys and the motivation behind this report.

DOC have requested a trend and power analysis of the braided river bird survey data that have been collected since 1986 on the Upper Rakitata to gain insight into the effectiveness of their monitoring protocol and to check the potential to use these data to investigate the impact of predator control on braided river bird populations in the area.

Objectives

We assessed the current and future potential, of the survey data from walk-through surveys of the Upper Rakitata River to detect river bird population changes. We focused on the ability (statistical power) to detect trends over time, with the eventual goal of being able to investigate the efficacy of predator control efforts in the area. We did this by:

1. Collating, cleaning and organising the data into a format fit for analysis of trends over time.
2. Determining the completeness of the dataset through time and identifying the most appropriate data to include in analyses.
3. Conducting a trend analysis on counts where consistent areas were surveyed for each species to see how braided river bird species populations are currently trending.
4. Running power analyses to quantify the ability of the current survey design to detect population changes in the count data after 20 years of continuous surveys.

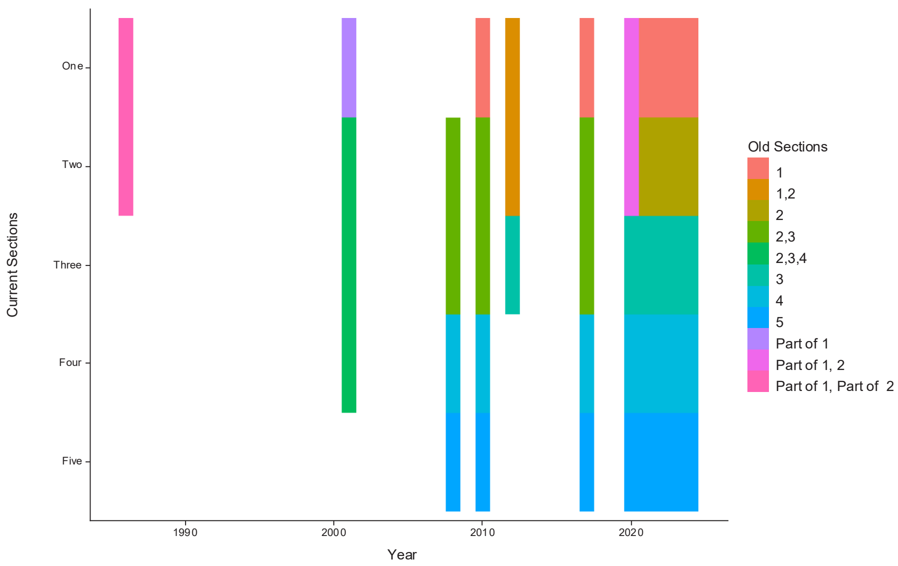


Figure 2 Years surveyed, and the sections recorded in the database. Old sections, represented by colours, are aligned with their current section equivalents on the y-axis.

Methods

Survey data

Data were collected by NZ Wildlife Service/DOC in the years 1986, 2001, 2008, 2010, 2012, 2017, and every year since 2020. Each year, over three days within the bird breeding season (October/November), between 8 and 20 observers walk downstream from the Clyde and Havelock confluence to the start of the Rakitata Gorge. The reach is currently broken into five sections, but the delimitation of reaches has varied over time (Figure 2). Counts of braided river bird species have been recorded by section. Due to the variation in the sections surveyed and the section naming/delimitation protocol, we have only included the most recent four years (2021–2024) of data in the following trend analyses. Counts from earlier years could potentially provide further information and allow longer trend evaluation but the issues mentioned above prevent robust analyses. We do, however, discuss challenges of longer-term analysis using historic data later in the report.

Discharge data

Discharge data for the Rakitata, collected at Klondyke, were provided by Environment Canterbury. These data were used to calculate days since last flood, defined as discharge above 250 m^3/s (pers. comm. Brad Edwards, DOC), and mean discharge across the survey. These variables were then tested as potential covariates in the trend analyses below.

Trend analysis

We modelled trends in bird counts over time using a generalised linear mixed modelling (GLMM) approach implemented using the lme4 package (Bates *et al*. 2015) in the R program (R Core Team 2022). River section was included as a random intercept to account for variability between sections. Retaining section-level data, rather than pooling counts across the river, preserves spatial information and avoids pseudo-replication. Modelling section as a random effect also partitions variance appropriately, providing more reliable estimates of year effects and greater power to detect population-level trends. An offset for the mean number of daily surveyors was included to account for variation in survey effort (mean number of daily surveyors) among years. We also considered including hectares surveyed as an additional offset, however, this metric was derived from polygons mapped in GIS using a single image across all years. Because that area measure would not reflect year-to-year changes in surveyed area and would be largely redundant once sections were treated as random intercepts, it was not retained in the final analyses.

Each species was modelled separately, with those having a median count of fewer than five individuals across years excluded due to data sparsity. Of the 14 species, this left seven for analysis. Mean days since the last flood (>250 m^3/s) was log-transformed, and continuous predictors (e.g. discharge and flood metrics) were centred and scaled to reduce collinearity and improve model convergence. Year was treated as a numeric variable coded 1–4 rather than scaled, to represent the sequential survey years directly.

We used small-sample-size corrected Akaike Information Criterion to compare alternative model structures, including either random or fixed slopes for each species model. Initial models were fit with a Poisson error distribution, however, after checking for overdispersion, all species were modelled using a Negative Binomial distribution. To evaluate the most parsimonious model for detecting temporal trends, we constructed candidate models that included the following predictors:

- Year
- Days since last flood (>250 m^3/s)
- Mean discharge across the survey (m^3/s)

An intercept-only model was also included for comparison. Candidate models were ranked using AICc, with ΔAICc and Akaike weights calculated using the *AICcmodavg* package (Mazerolle 2023) in R. ΔAICc values provide a measure of relative support, with values close to zero indicating strong support. As a general guideline, models within 2 units of the best model have substantial support, values of 4–7 indicate considerably less support, and values >10 indicate essentially no support (Burnham & Anderson 2002). Akaike weights provide a normalised measure of relative support across the candidate set and can be interpreted as the probability that a given model is the best of those considered for explaining the data. In this framework, a “better” model is one with lower AICc, smaller ΔAICc, and higher Akaike weight, indicating stronger empirical support relative to alternatives. We also calculated marginal and conditional R^2 values using the *r.squaredGLMM* function in the R package MuMIn (Bartoń 2010), with marginal R^2 representing the variance explained by fixed effects alone and conditional R^2 representing the variance explained by both fixed and random effects.

For formal testing of trends in the 2021–2024 dataset, we fit a standardised model for all species of the *form: glmm.nb(count ~ year + (1 | section_number), offset = log(mean_daily_surveyors))*. Significance of the Year effect was assessed using likelihood ratio tests (LRT). These same models formed the basis for subsequent power analyses.

The number of surveyors changed only slightly over the four years (ranging from 12 to 15 surveyors), although there was a slight increase through time (Figure S1A). To assess the sensitivity of our results to the surveyor offset, we repeated the same models without including the offset term, tested for temporal trends, and also used these alternative models in the power analyses. Results from these models are presented in the supplementary material and are discussed in the section *Comparison of offset and no-offset models*.

Power analysis

For each species model, we used the R package *simr* (Green & MacLeod 2016) to simulate larger sample sizes by extending the number of time points and applying a range of effect sizes. Models were extended to include an additional 1 to 16 years surveyed (a total of 20 surveys) under five scenarios of population change (-10%, -5%, 0%, 5% and 10%) per year, resulting in 80 model combinations. For each extended model, we ran 999 simulations using the "lr" (likelihood ratio test) method to calculate the statistical power to detect the specified effect size. A power threshold of ≥80% was considered to indicate adequate ability to detect a given trend.

All analyses were conducted using R 4.4.3 (R Core Team 2022). Graphs were produced using ggplot2 (Wickham 2016) and ggeffects (Lüdecke 2018) was used to extract marginal means and 95% confidence bands from the mixed models.

Results

Trend analysis / test models

After selecting random effects for each species, all models included a fixed slope with intercepts varying by section (Table 1). For banded dotterel, black-fronted tern, black-billed gull, South Island pied oystercatcher, and spur-winged plover, the intercept-only model was consistently top-ranked, indicating that none of the tested predictors provided additional explanatory power (Table 2; Figure 1). For southern black-backed gull and wrybill, models including Year as the sole fixed effect ranked highest, while for black-billed gull a model with Year + Mean Discharge Across Survey was selected (Table 2). However, this model performed only marginally better than the Year-only model, and further inspection showed strong collinearity between Year and Mean Discharge (Figure S2). Therefore, Year-only models were used for trend analysis.

Across all species, marginal R^2 values were very low, while conditional R^2 values were comparatively higher (Table 2). This pattern indicates that the fixed effects of year and mean discharge explained little variation in bird counts, whereas differences among sections accounted for most of the explained variation. This means that temporal trends at the scale of the whole river are challenging to detect with the current dataset, because section-level variation dominates. While it is possible that strong changes within individual sections could still be identified, the overall statistical power to detect consistent year-to-year changes across the river is low.

Despite most species showing no significant trends, three species did show statistically significant changes in counts over the 2021–2024 period (Table 3, Figure 3). Southern black-backed gull counts declined by ~40% per year (95% CI: -53% to -24%), a pattern likely reflecting ongoing management efforts to control this species to safeguard endangered braided river bird species. Wrybill numbers also declined, with a smaller but still significant decrease of ~7% per year (95% CI: -11% to -2%). In contrast, black-billed gull counts increased by ~105% per year (95% CI: 21% to 232%). However, this result should be treated with caution, because it was strongly influenced by a large outlier in Section 2 in 2023: we encourage double checking for a possible data entry error here. Overall, while these results highlight some trends, it is important to note that the fixed effects in the models generally performed poorly and explained little of the variation in bird counts (Table 2).

Power analyses

The power to detect a ±10% annual change with ≥80% power was variable between species (Figure 4). However, analyses from most species indicated approximately 10–15 years of data would be sufficient to detect a ±10% annual chang with ≥80% power (Figure 4). A notable exception was black-billed Gull, where a -10% annual change would require almost 20 years of data. Detecting smaller effects (±5% per year) generally required additional survey years. Analyses indicated wrybill, banded dotterel and South Island pied oystercatcher needed between 10–15 years of monitoring to reach 80%

power to detect a ±5% annual change, whereas black-fronted tern required 18–20 years and neither black-billed gull nor spur-winged plover achieved this threshold within the 20-year window (Figure 4). These results indicate that more substantial population changes may be detected within a decade, but identifying more subtle shifts will require much longer monitoring from sustained (i.e., long-term) data collection.

Comparison of offset and no-offset models

We compared models with and without an offset for survey effort (log of surveyor-numbers). Results differed in both effect sizes and apparent significance (Table S1, Figure S2). In the no-offset models, slopes tended to be larger and power analyses suggested detectability of year effects was higher. However, we are cautious about this result because it likely reflects the confounding influence of increasing survey effort over time (Figure S1B). We are concerned that more observers would result in more birds being recorded and that this may mimic population growth if not accounted for. In contrast, the offset models specifically account for effort and therefore provides estimates of change in bird abundance per unit effort. As expected, the offset models produced more conservative slopes and lower apparent power, but also higher marginal fixed effects in many cased (Table S1, Figure S3).

Challenges with analysing trends over the full dataset

Survey effort has varied among years, both in the areas of river surveyed (Figure 2) and in the number of surveyors. In addition, counts from some sections have been combined, while others have been only partially named, creating inconsistencies that limit robust analysis of long-term trends with the earlier (pre 2021) count data. To enable meaningful year-to-year comparisons, survey effort needs to be standardised in terms of both area covered and likely surveyor numbers. Sections labelled as “part” sections are particularly problematic, because there is no information on what proportion of the currently defined section was surveyed. Without further clarification of survey boundaries in these cases, we recommend excluding them from any analyses. The inconsistent lumping of other sections across years further complicates modelling approaches. For example, river section cannot be used in a random effect because the spatial units are not consistently defined through time. This is the major issue preventing the older counts from being used in a robust way.

To provide at least a preliminary view of longer-term patterns, we selected the most consistently surveyed sections (2, 3, and 4), summed counts within these sections for each year they were surveyed, and plotted these data for visual inspection (Figure 5). We did not do any formal analyses of these data, because the amount of information available is insufficient to yield statistically robust results.

Reflections and recommendations

Overall, our analyses indicate there is potential to detect trends of ±5% change in braided river bird counts for some of the braided river species with ~80% certainty if monitoring is continued over the next decade. However, to evaluate the efficacy of predator control, baseline data from periods without predator management are essential. Given the lack of sufficient pre-management data for the Upper Rakitata, it will be challenging to robustly attribute changes in bird populations to predator management with the current dataset. Below we outline key reflections on the current dataset and recommendations for future monitoring.

1. Clarify goals

We encourage clarification of the goals of monitoring and the types of management actions that walk-through survey data are intended to inform. This would include determining the effect size magnitude (% population change) that would be desirable to detect for any management action. This prioritisation is especially important if assessing predator control outcomes is a priority, because the current dataset alone is unlikely to support such an analysis, especially over the short term (< 10 years).

2. Data integrity

There appear to be potential data entry errors (e.g., 2023 black-billed gull counts in Section 2, and 2021 Southern black-backed gull counts in Section 5). We recommend double-checking these records to ensure data quality control. The number of surveyors is also missing from the 2001 dataset.

3. Data collection and covariates

A key limitation of the current dataset is the large amount of unexplained variation and the lack of consistently recorded covariates (described below). We recommend reviewing the data flow from field sheets through to the central database so that useful metadata are captured routinely. The Environment Canterbury (ECan) database is currently very tidy and well structured, which is a strength, but it is missing potentially valuable information such as the number of surveyors per section, their experience level, survey duration, and weather conditions. Additionally, if the number of birds an individual surveyor counts are recorded in the field, along with where in the river they walked, then keeping this individual level data would also help with future modelling of observer effort and variation.

We also encourage future investigation of potential covariates that might account for some of the unexplained variation, such as habitat or area-based measures (e.g., aquatic–terrestrial edge length), which could be retrospectively derived from satellite imagery. While it is uncertain how much additional explanatory variables will ultimately improve model performance, systematically recording data collected during the surveys and storing them would preserve the option to test their utility in future analyses. Moreover, if an additional covariate that explained a large amount of habitat-based variation in bird abundance could be found, then it could be a game changer by greatly increasing statistical power. The general goal in the search for additional covariates should be to statistically account for local variation available bird habitat (with the covariate[s]) so that variation due to other causes can be detected.

4. Measures of effort

To compare data across years, it is essential that either survey effort is kept consistent or that counts are standardised for effort. In our analyses we tested both section area and number of surveyors as potential measures of effort. Section area was not suitable because it was static across years, whereas the number of surveyors provided a more useful offset allowing modelling of per-surveyor counts. Including this offset sometimes altered the estimated trends, highlighting its importance and the sensitivity of trends to its inclusion. As above, we reiterate our recommendation to consistently record in the database those variables related to survey effort.

The models that included a measure of surveyor effort currently provide the most robust basis for inference. However, an important future step will be to better quantify how survey effort translates into bird detectability. The current offset assumes a linear 1:1 relationship between the number of surveyors and the number of birds detected/counted, but this is unlikely to hold in practice. Two potential future approaches include repeat-survey methods or double-observer methods. For example, a section could be surveyed multiple times in close succession with different team sizes allowing the estimate of changes in detection probability changes across survey effort. Another possible option could be to have two observer teams survey the same section simultaneously but independently, this would enable estimation of the proportion of birds missed. It is possible that even limited implementation of these approaches (e.g., in selected sections) could provide a stronger empirical basis for the offset, reveal whether additional surveyors yield diminishing returns, and allow the standardisation of long-term trends with greater confidence. Therefore, we recommend a future review and investigation of additional survey methods for improving detection probabilities that could implemented in the Upper Rakitata walk-through surveys.

5. Alternative survey designs

If the goal is to detect broad-scale trends, all river sections may not need to be monitored every year. Monitoring a subset of key sections could free up resources to intensify effort within them, improving data quality by resolving some of the issues mentioned above. Adjusting survey frequency is another possible option. For example, a 3-years-on / 3-years-off design has been proposed as an alternative that would reduce costs. However, this would likely substantially reduce statistical power and extend the time required to detect changes. Less disruptive alternatives include reducing the number of sections surveyed annually or possibly adopting a 2-years-on / 1-year-off cycle. Importantly, understanding the inter-annual variation in bird counts is crucial for making informed decisions about how reducing the frequency of surveys will impact power to detect trends. For example, if there is high interannual variability then increasing the frequency of surveys might miss important trends and we strongly recommend further analyses to explore these scenarios before changes are adopted.

We also encourage, depending on the goals of the braided bird monitoring program, considering if there are other complementary methods (e.g. close-kin mark–recapture, and drone-based aerial surveys) which could be considered and trialled, and which may improve efficiency and robustness as technology develops.

6. Expectations and null models

An additional approach to consider involves 'null models'. Here, the population changes expected if predator control were effective, based on each species' biology, reproductive rate, and likely trajectory without control could be developed. This would provide a 'null model' against which walk-through data could be compared. Importantly, predator control may stabilise populations rather than increase them. In such cases, detecting "no change" should still be interpreted as a positive outcome, since it implies that predator control has prevented an ongoing decline.

It could be possible to build species-specific population models or to draw on published estimates of natural decline rates under predation pressure. These counterfactual models could then be used to benchmark observed results. For example, if a species is expected to decline by ~5% annually without predator control, then a flat or stable trend in counts under predator management would represent a clear management benefit. Such an approach would shift the interpretation of results from "did predator control increase populations?" to "did predator control prevent expected declines?" Such an approach could provide a more feasible way to gain causal inference in a situation where decades of data are needed to evaluate a management regime.

7. Detecting predator control impacts

Currently, the survey data provide no ability to attribute population changes directly to predator control, so different approaches should be considered. For example, to directly assess the impacts of predator control, experimental approaches could be valuable. A possible experimental design could include intensifying control in some sections but not others and directly comparing outcomes. Another approach for understanding impacts of predator control would be to investigate cross-catchment analyses (e.g., using the ECan database) which may provide useful contrasts. For example, there is a chance that even if predator control is successful, populations could decline but that decline is not attributed to predators. Assigning causality will be challenging but there would be more strength of inference if there were cross-catchment comparisons. While these comparisons will introduce additional noise, the accessibility and breadth of the data in the current database make them worth investigating. Any such analyses will require careful thought about how predator control effort (or success) is quantified between catchments, including trap effort, target species, and residual predator population estimates.

Conclusion

The current dataset provides a valuable foundation for future surveys but has clear limitations that constrain the detection of subtle trends and means evaluation of predator management impacts requires long-term (10+ years) surveys. Continued collection of braided river bird monitoring data will be essential to detect population changes, but it should be accompanied by careful reflection on survey goals, including what magnitude of change is considered meaningful (or necessary) for management. Larger changes are easier to detect but may represent declines that are already too late to reverse, so prioritising early detection of smaller but meaningful changes appears the best approach. At the same time, monitoring methods, metadata capture, and survey design should continue to be refined, and complementary approaches explored. Most importantly, to ensure the braided river walk-through survey continues to be fit-for-purpose, we encourage monitoring to be an iterative process of data collection, evaluation, and improvement. This involves critical evaluation of goals, data and methods that will maximise the long-term value of this programme and its ability to inform management decisions.

Table 1. Random effects model selection for each species using small-sample-size corrected Akaike Information Criterion (AICc). Models compared included intercepts varying by section (1 | Section) and slopes for Year varying by section (Year | Section). Reported are AICc, ΔAICc, and AICc weights (Weight).

Species	Random effect	AICc	ΔAICc	Weight
Banded dotterel	(1 | Section)	225.50	0.00	0.98
	(Years | Section)	232.89	7.39	0.02
Black-fronted tern	(1 | Section)	243.64	0.00	0.95
	(Years | Section)	249.45	5.81	0.05
Black-billed gull	(1 | Section)	174.92	0.00	0.98
	(Years | Section)	182.60	7.68	0.02
Southern black-backed gull	(1 | Section)	237.47	0.00	0.98
	(Years | Section)	245.14	7.67	0.02
Spur-winged plover	(1 | Section)	127.57	0.00	0.98
	(Years | Section)	135.36	7.79	0.02
South Island pied oystercatcher	(1 | Section)	193.01	0.00	0.98
	(Years | Section)	200.75	7.73	0.02
Wrybill	(1 | Section)	170.87	0.00	0.98
	(Years | Section)	178.35	7.48	0.02

Table 2. Model selection for fixed effects of Year, Mean Days Since Last Flood, Mean Discharge Across Survey (labelled Flow), and their interactions, compared to an intercept-only model. Shown are Akaike Information Criterion (AICc), ΔAICc, AICc weights (Weight), and marginal (mR^2) and conditional (cR^2) R^2 values. Likelihood ratio test results for Year-only models are provided in Table 3.

Species	Fixed effects	AICc	ΔAICc	Weight	mR^2	cR^2
Banded dotterel	Intercept only	224.149	0.000	0.561	0.000	0.308
	Years	225.498	1.349	0.286	0.037	0.369
	Years + Mean Days Since Flood	228.158	4.009	0.076	0.055	0.397
	Years + Mean Flow Across Survey	228.635	4.486	0.060	0.046	0.389
	Years * Mean Flow Across Survey	232.316	8.166	0.009	0.056	0.399
	Years * Mean Days Since Flood	232.316	8.166	0.009	0.056	0.399
Black-fronted tern	Intercept only	240.947	0.000	0.708	0.000	0.096
	Years	243.636	2.689	0.185	0.034	0.039
	Years + Mean Flow Across Survey	245.985	5.038	0.057	0.085	0.085
	Years + Mean Days Since Flood	246.916	5.970	0.036	0.044	0.044
	Years * Mean Flow Across Survey	250.132	9.185	0.007	0.086	0.086
	Years * Mean Days Since Flood	250.132	9.185	0.007	0.086	0.086
Black-billed gull	Years + Mean Flow Across Survey	173.942	0.000	0.397	0.245	0.736
	Years	174.920	0.978	0.243	0.146	0.681
	Intercept only	176.071	2.129	0.137	0.000	0.527
	Years * Mean Flow Across Survey	176.906	2.964	0.090	0.269	0.758
	Years * Mean Days Since Flood	176.906	2.964	0.090	0.269	0.758
	Years + Mean Days Since Flood	178.365	4.423	0.043	0.156	0.680
Southern black-backed gull	Years	237.469	0.000	0.698	0.300	0.651
	Years + Mean Flow Across Survey	240.857	3.388	0.128	0.306	0.653
	Years + Mean Days Since Flood	240.975	3.505	0.121	0.302	0.656
	Years * Mean Flow Across Survey	244.237	6.768	0.024	0.319	0.676
	Years * Mean Days Since Flood	244.237	6.768	0.024	0.319	0.676
	Intercept only	247.296	9.827	0.005	0.000	0.256
South Island pied oyster-catcher	Intercept only	189.859	0.000	0.772	0.000	0.253
	Years	193.014	3.155	0.159	0.000	0.253
	Years + Mean Flow Across Survey	196.144	6.286	0.033	0.007	0.260
	Years + Mean Days Since Flood	196.547	6.688	0.027	0.001	0.253
	Years * Mean Flow Across Survey	200.292	10.433	0.004	0.007	0.261
	Years * Mean Days Since Flood	200.292	10.433	0.004	0.007	0.261
Spur-winged plover	Intercept only	125.451	0.000	0.630	0.000	0.115
	Years	127.570	2.118	0.218	0.010	0.128
	Years + Mean Flow Across Survey	129.529	4.077	0.082	0.027	0.148
	Years + Mean Days Since Flood	131.189	5.737	0.036	0.010	0.128
	Years * Mean Days Since Flood	132.676	7.225	0.017	0.033	0.159
	Years * Mean Flow Across Survey	132.676	7.225	0.017	0.033	0.159
Wrybill	Years	170.874	0.000	0.457	0.017	0.349
	Intercept only	171.387	0.513	0.353	0.000	0.329
	Years + Mean Days Since Flood	174.172	3.298	0.088	0.018	0.351
	Years + Mean Flow Across Survey	174.359	3.485	0.080	0.017	0.350
	Years * Mean Days Since Flood	178.347	7.473	0.011	0.018	0.351
	Years * Mean Flow Across Survey	178.347	7.473	0.011	0.018	0.351

Table 3. Results of likelihood ratio tests (LRT) for temporal trends in per-surveyor counts of river birds from 2021–2024. Models were fitted using a negative binomial distribution with Years as a fixed effect, section as a random intercept, and an offset for mean daily surveyors. Reported are the test statistic (χ^2), associated p-value, slope estimates, 95% confidence intervals, and the corresponding annual percentage change in per-surveyor counts. Significant trends ($p < 0.05$) are bolded. Marginal and conditional R^2 values for these models are presented in Table 2.

Species	Fixed effects	LRT (c^2)	LRT *p*	Slope	Lower CI	Upper CI	Annual % change (per surveyor)
Banded dotterel	Years	1.947	0.163	-0.099	-0.230	0.049	
Black-fronted Tern	Years	0.506	0.477	-0.137	-0.458	0.183	
Black-billed gull	Years	4.801	**0.028**	0.720	0.193	1.202	Per-surveyor counts **increased** by **~105% per year** (95% CI: 21% to 232%)
Southern black-backed gull	Years	19.642	**<0.005**	-0.508	-0.752	-0.275	Per-surveyor counts **declined** by **~40% per year** (95% CI: −53% to −24%)
South Island pied oyster-catcher	Years	0.011	0.916	0.007	-0.117	0.116	
Spur-winged Plover	Years	1.084	0.298	0.130	-0.104	0.382	
Wrybill	Years	4.303	**0.038**	-0.067	-0.126	-0.008	Per-surveyor counts **declined** by **~7% per year** (95% CI: −11% to −2%)

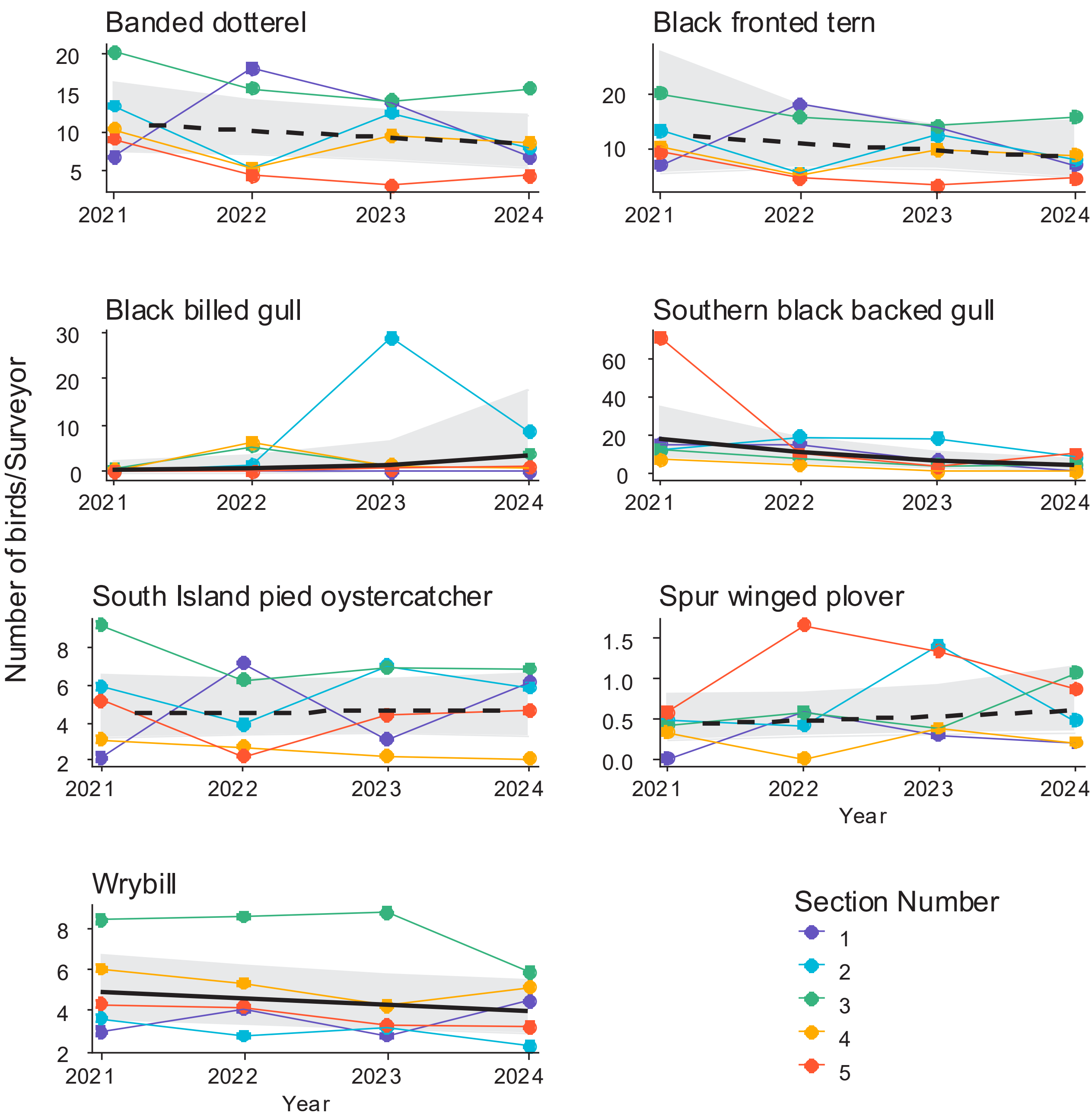


Figure 3. Fitted generalised linear mixed effects models (GLMMs) showing temporal trends in per-surveyor counts of braided river bird species from 2021–2024. Each panel represents one of the seven bird species, with coloured lines indicating counts by section (random intercept). Grey shading shows 95% confidence intervals around the fitted relationship. Solid black lines indicate significant temporal trends based on likelihood ratio tests (LRTs), while dashed black lines indicate non-significant trends. See Tables 2 and 3 for corresponding model selection results and test statistics.

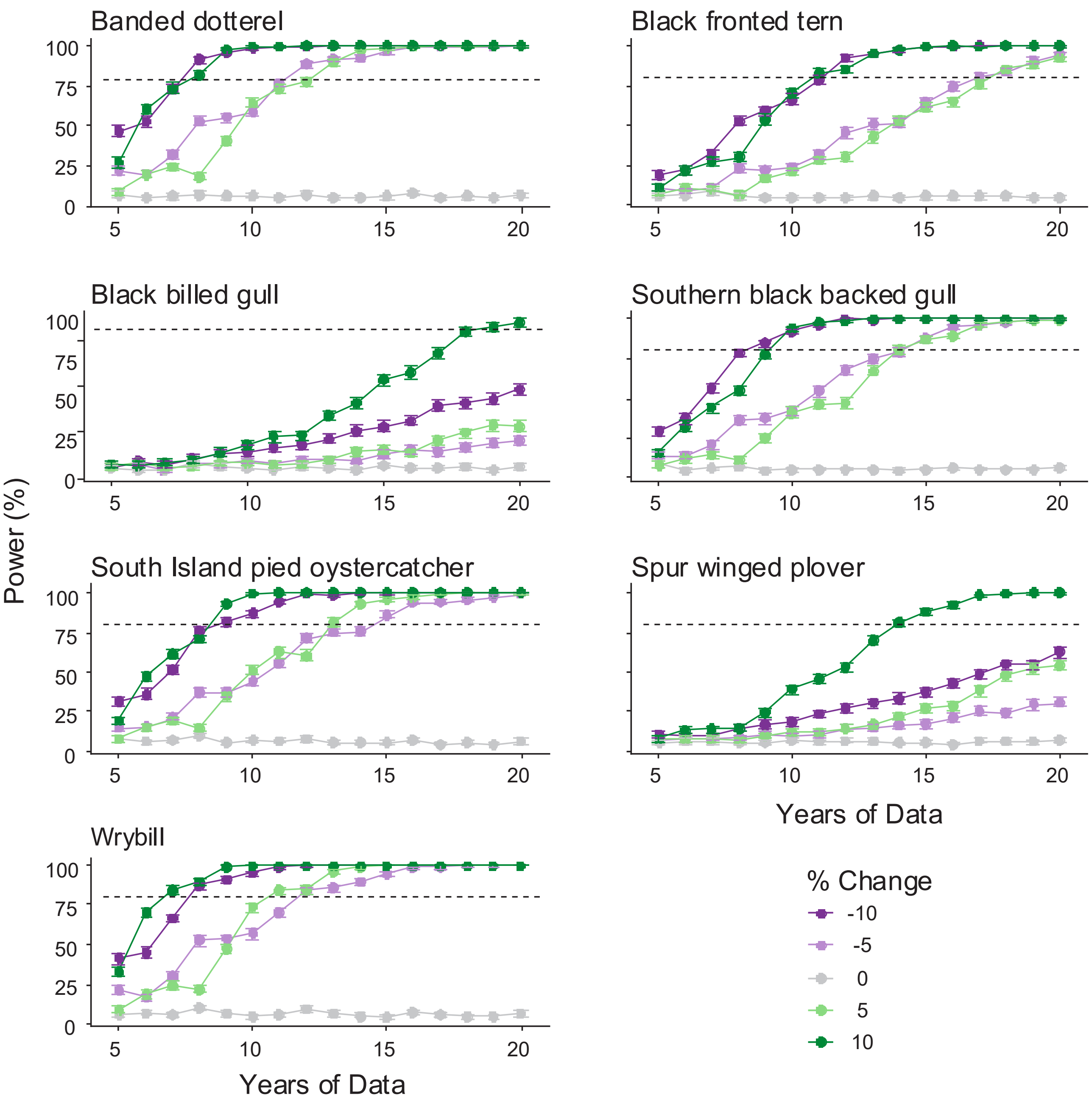


Figure 4. Results of power analyses for detecting temporal trends in per-surveyor counts of braided river bird species. Each panel shows one of the seven bird species, with lines representing different annual effect sizes (-10%–10% change per year). Models were extended to represent 5–20 years of data, with 999 simulations per model using the “lr” method to calculate statistical power. Error bars show 95% confidence intervals around the estimated power. The horizontal dashed line indicates the 80% power threshold commonly used to define adequate ability to detect a trend.

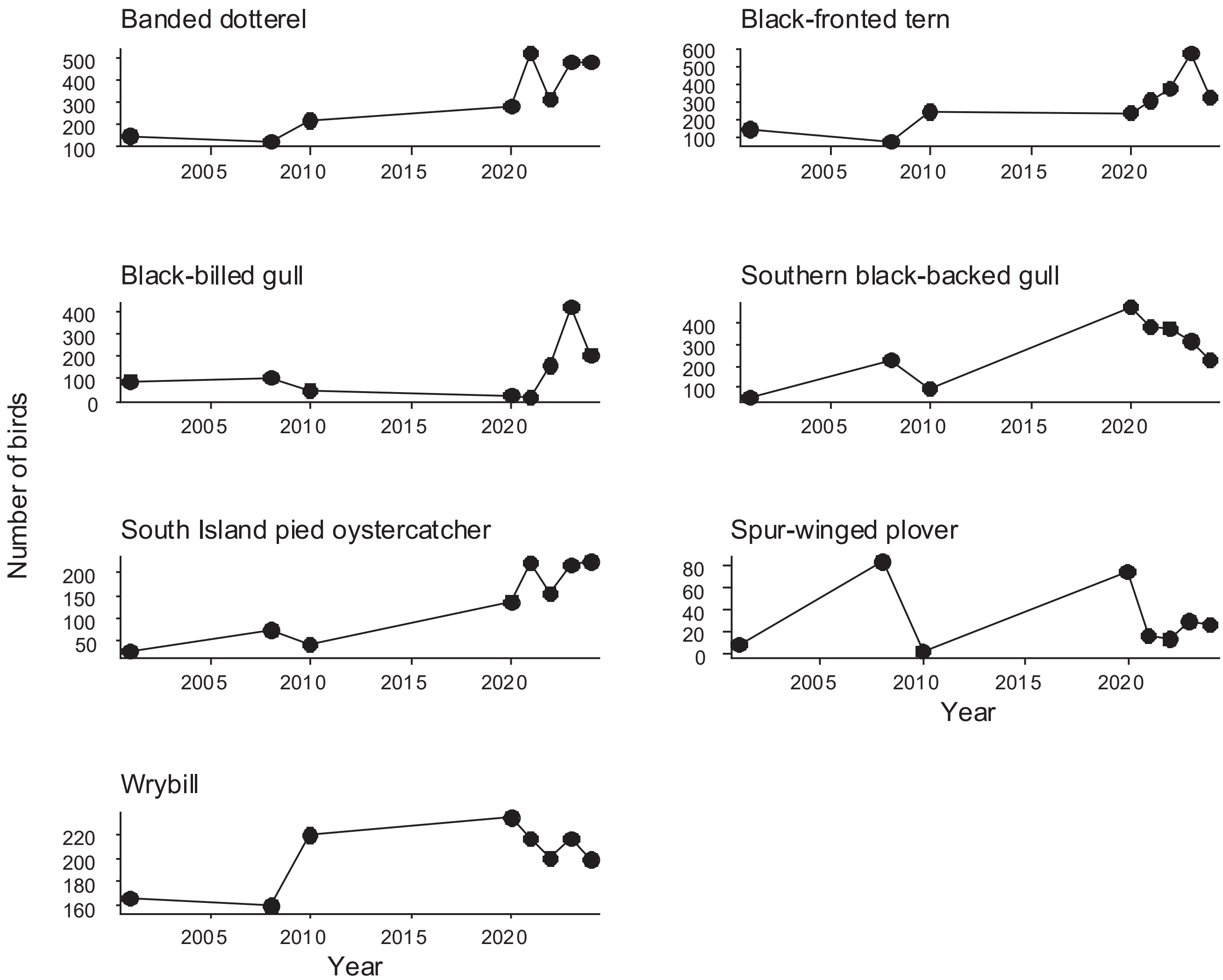


Figure 5. Black points indicate annual bird totals, calculated by summing counts from sections 2, 3, and 4, which were the most consistently monitored across survey years. These data are presented for visual inspection only and were not formally analysed due to inconsistencies in survey effort, section definitions, and observer numbers across years. The plots provide only a preliminary indication of possible long-term patterns and should not be interpreted as robust trend analyses.

References

Bartoń K 2010. MuMIn: Multi-Model Inference. R package version 1.48.4.

Bates D, Mächler M, Bolker B, Walker S 2015. Fitting Linear Mixed-Effects Models Using lme4. Journal of Statistical Software 67(1).

Burnham KP, Anderson DR 2002. Model selection and multimodel inference: a practical information-theoretic approach. Springer.

Green P, MacLeod CJ 2016. SIMR: An R package for power analysis of generalized linear mixed models by simulation. Methods in Ecology and Evolution 7(4): 493–498.

Lüdecke D 2018. ggeffects: Tidy Data Frames of Marginal Effects from Regression Models. Journal of Open Source Software 3(26): 772.

Mazerolle MJ 2023. AICcmodavg: model selection and multimodel inference based on (Q) AIC (c). R package version 2.3-2. 2023.

O'Donnell CFJ, Sanders M, Woolmore C, Maloney RF 2016. Management and research priorities for conserving biodiversity on New Zealand's braided rivers.

R Core Team 2022. R: A Language and Environment for Statistical Computing. Vienna, Austria, R Foundation for Statistical Computing.

Wickham H 2016. ggplot2: elegant graphics for data analysis. Second edition. Springer international publishing.

Supplementary material

Table S1. Llikelihood ratio tests (LRT) for temporal trends in of river birds from 2021–2024. Models were fitted using a negative binomial distribution with Year as a fixed effect, section as a random intercept, and no offset. Reported are the test statistic (χ^2), associated p-value, slope estimates, 95% confidence intervals, marginal and conditional R^2 values.

Species	LRT (c^2)	LRT p	Slope	Lower CI	Upper CI	mR^2	cR^2
Banded dotterel	0.098	0.755	-0.022	-0.184	0.121	0.003	0.459
Black-fronted tern	0.098	0.755	-0.022	-0.150	0.174	0.003	0.459
Black-billed gull	6.238	0.013	0.800	0.202	1.306	0.186	0.733
Southern black-backed gull	14.185	<0.005	-0.429	-0.629	-0.239	0.252	0.665
South Island pied oyster-catcher	1.690	0.194	0.084	-0.058	0.215	0.043	0.498
Spur-winged plover	2.973	0.085	0.209	-0.079	0.426	0.090	0.503
Wrybill	0.114	0.735	0.011	-0.055	0.084	0.001	0.820

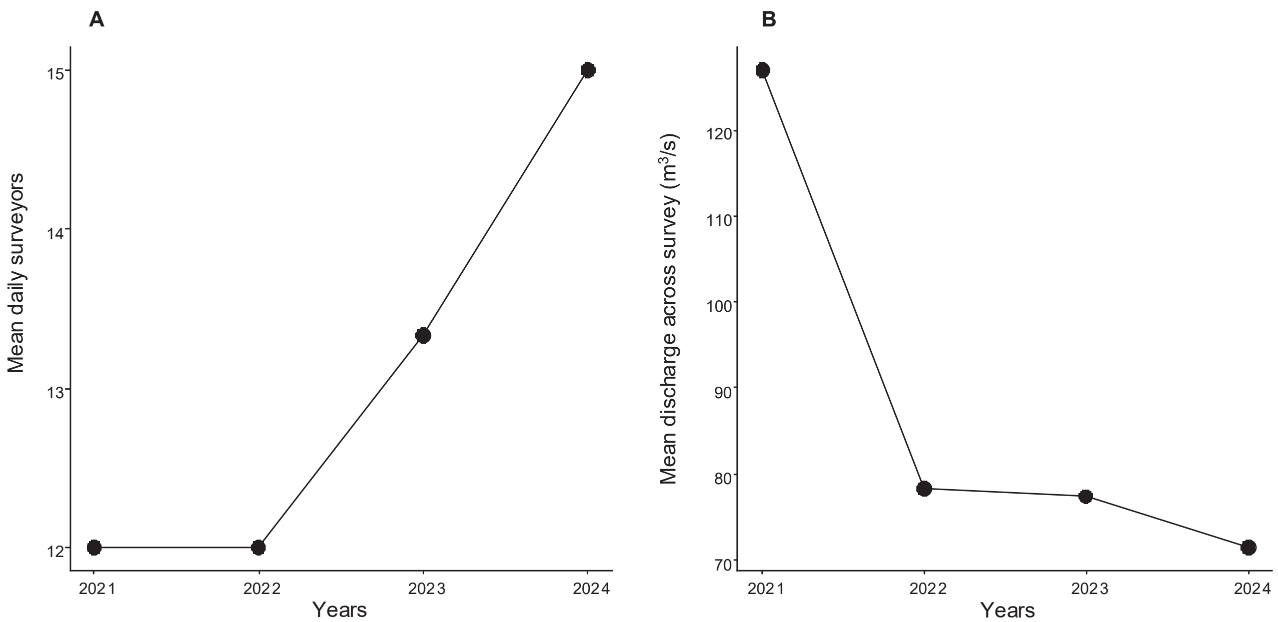


Figure S1. Mean number of daily surveyors involved in braided river bird counts (A) and annual variation in mean river discharge (m^3/s) across bird survey periods (B)

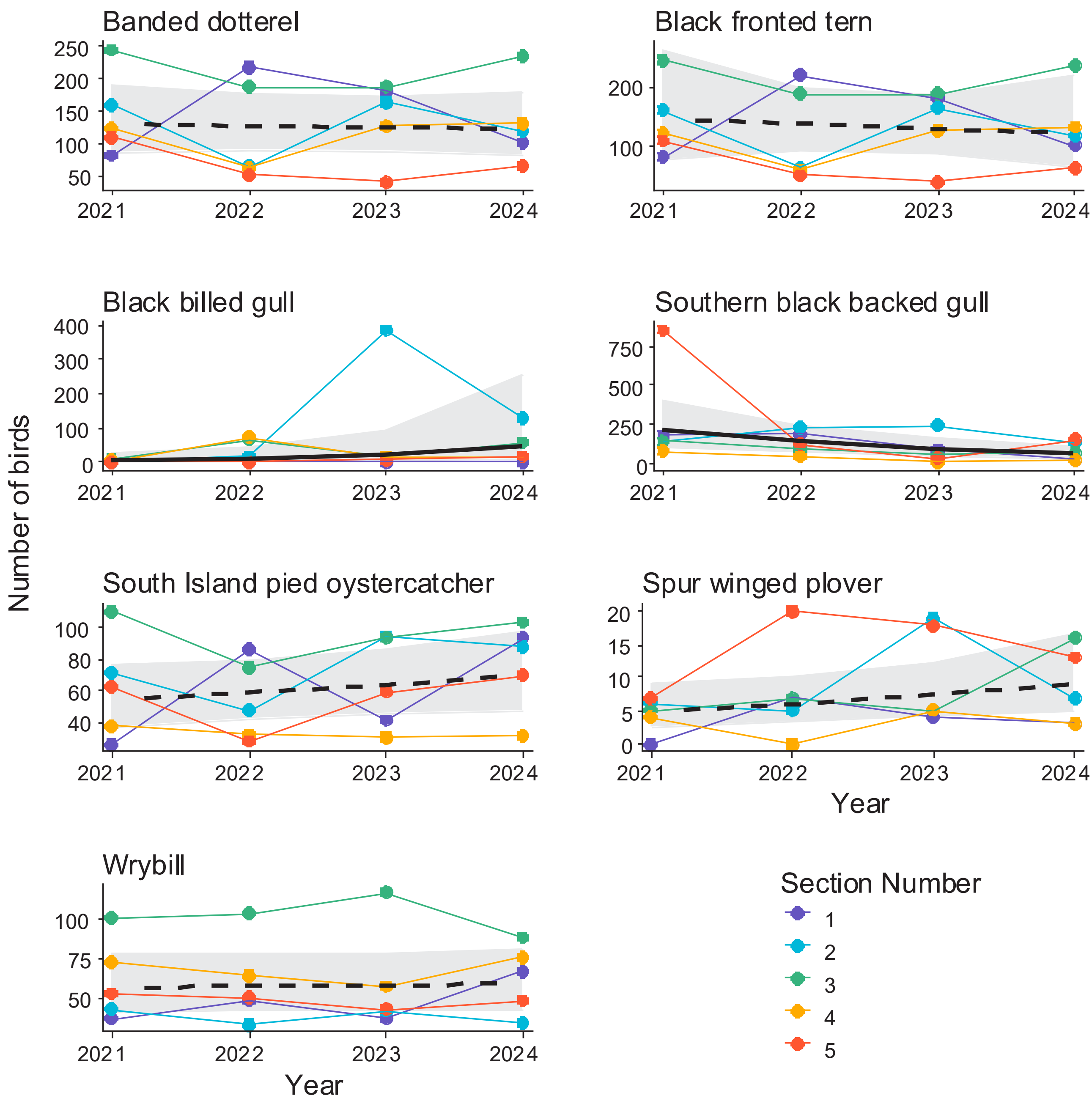


Figure S2. Fitted generalised linear mixed effects models (GLMMs) showing temporal trends in counts of braided river bird species from 2021–2024, using models fitted with no offset. Each panel represents one of the seven focal species, with coloured lines indicating counts by section (random intercept). Grey shading shows 95% confidence intervals around the fitted relationship. Solid black lines indicate significant temporal trends based on likelihood ratio tests (LRTs), while dashed black lines indicate non-significant trends. See Tables 2 and 3 for corresponding model selection results and test statistics.

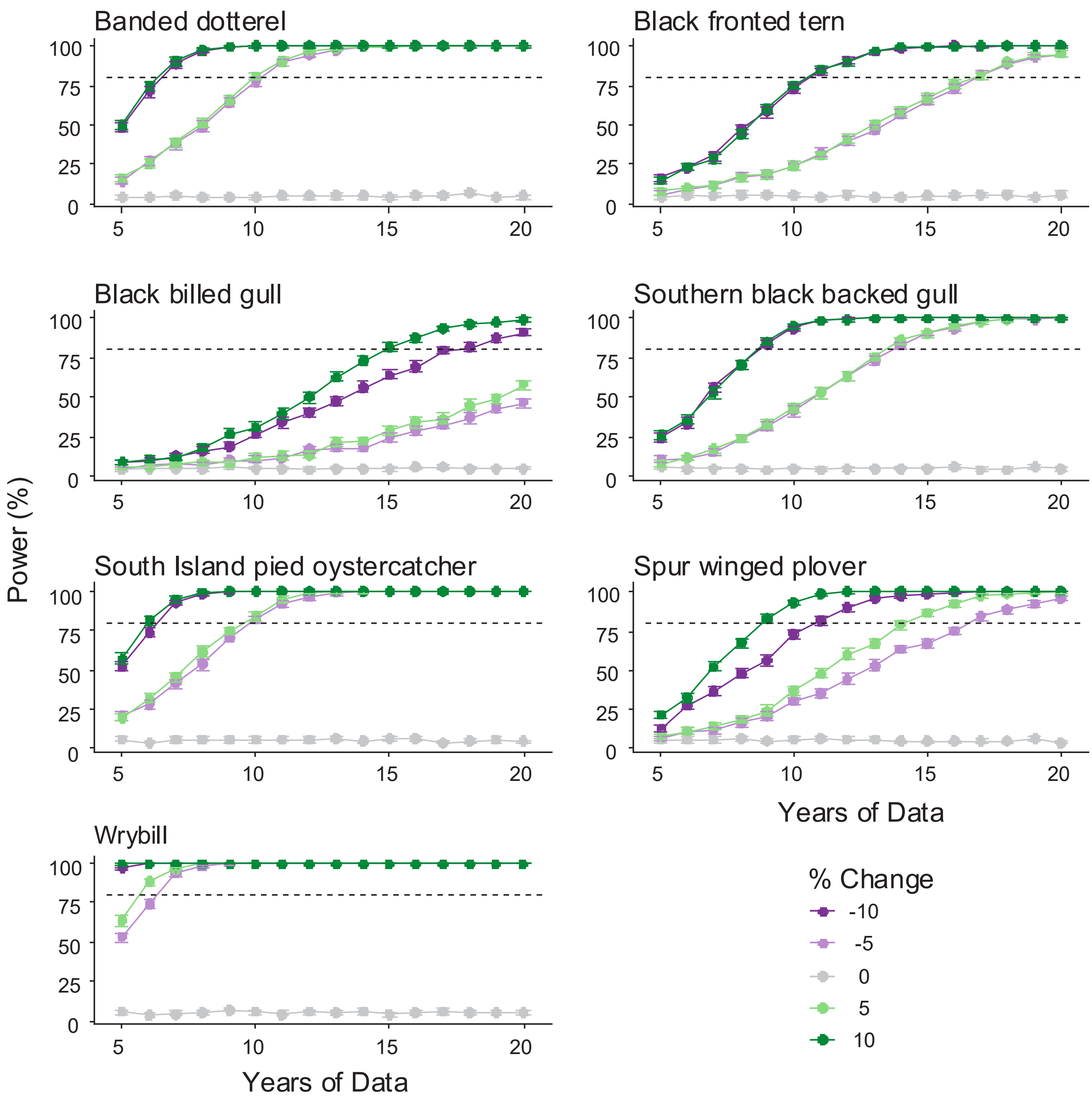


Figure S3. Results of power analyses for detecting temporal trends in counts of braided river bird species, using models with no offset. Each panel shows one of the seven focal species, with lines representing different annual effect sizes (-10%–10% change per year). Models were extended to represent 5–20 years of data, with 999 simulations per model using the "lr" method to calculate statistical power. Error bars show 95% confidence intervals around the estimated power. The horizontal dashed line indicates the 80% power threshold commonly used to define adequate ability to detect a trend

School of Biological Sciences
Te Kura Pūtaiao Koiora

T: +64 3 369 5200
E: biology@canterbury.ac.nz

University of Canterbury
Te Whare Wānanga o Waitaha
Private Bag 4800
Christchurch 8140
New Zealand

www.canterbury.ac.nz/science/biology