# Data analysis to support monitoring of the Rivers of Carbon and Boorowa River Recovery projects

Robin Hale[1] and Paul Reich

[1]School of BioSciences, University of Melbourne

## 1. Executive Summary

The Boorowa River Recovery (BRR) and Rivers of Carbon (RoC) projects were established to improve the condition of riparian zones in the upper Murrumbidgee and Lachlan catchments. As part of these projects, monitoring has been conducted and a large amount of data were collected based on a sampling design developed by personnel from CSIRO. Here we analyse this data to (1.) evaluate responses to riparian works and (2.) guide future sampling.

We highlight the following key messages and recommendations from our analyses of data from these two projects:

(1.) The BRR project has been monitored biennially from 2008-2014. In many cases, expected short-term (after six years) responses to works were detected (e.g. increased shrub cover, incidence of native flowering, and leaf litter, decreased bare ground).

(2.) BRR sampling was initially undertaken using both cross-section and transect methods. However, the sampling effort used for the cross-section method (one cross-section per site) was found to be inadequate for assessing changes in structural vegetation within the riparian zone. Cross-section sampling was phased out in 2014.

(3.) Continued biennial monitoring of the BRR project is unlikely to be warranted, given that further responses to riparian works are likely to occur over longer-time frames. We recommend that sites be resampled in 5-6 years (i.e. roughly a decade after works) to assess medium to long term changes. In the meantime, it is important that routine inspections are undertaken at sites to ensure the compliance of works (e.g. are plants alive, are fences intact) and to maintain relationships with landholders.

(4.) Monitoring at works sites may show a change has occurred but control sites are necessary to distinguish responses to works from temporal changes caused by other factors. It is therefore vital that any monitoring includes both control and works sites. We recommend that some of the BRR control sites are included in future RoC sampling to provide controls for RoC works sites. Ideally a minimum of 5-6 of these proxy controls should be sampled to properly describe background variability.

(5.) The BRR results can be used to guide the frequency of sampling in the RoC project. Ideally RoC sites should be sampled in the year prior to the implementation of works to assess pre-works condition, then again in the year immediately following works. Re-sampling sites after three and five years would allow short-term responses to be assessed. The overall aim of the RoC project, to link remnant vegetation and previously rehabilitated sites to form intact riparian corridors, will require monitoring for longer than the initial six year lifespan of the project. As with the BRR project, it is important that routine inspections of sites are undertaken in years no actual sampling is to be completed.

(6.) It is important to consider the sensitivity of the indicators being monitored when developing expectations about the rates of change in riparian characteristics. While a number of short-term changes may occur immediately following works, overall changes in RARC scores will likely take several years to occur. Relative changes in individual indicators must be relatively large for scores to change. For example, an increase in canopy cover from ~20 to >60% cover represents a shift from a score of 1 to 3 but this may take 15-20 years to occur. Overall increases in RARC scores will also be dependent on increases in a range of different sub-indices and indicators.

(7.) While the RARC method has a number of logistical advantages (e.g. minimal technical expertise is required, speed of monitoring each site), more subtle responses to works could be characterised by monitoring a smaller selection of key indicators in greater detail i.e. using more sensitive and repeatable methods. These indicators could be selected to encompass the key changes in structural vegetation (e.g. bare ground, leaf litter, shrub cover, canopy cover).

(8.) The assumptions that we have made in our analyses here can be updated as more monitoring data are gathered. For example, power analyses can be re-run with updated estimates of background variability in the absence of works, or to assess the potential detectability of non-linear responses.

# 2. Introduction

Riparian zones act as the interface between aquatic and terrestrial ecosystems, and are areas of exceptionally high productivity and biodiversity. Riparian zones fulfil a number of key ecological roles integral to the function of both aquatic and terrestrial ecosystems, for example moderating stream temperature, mediating inputs of sediment and nutrients into streams, providing habitat for animals, contributing to food webs, and acting as dispersal corridors and refuges for plants and animals.

In many areas of Australia, riparian zones have been extensively degraded, and many of the key ecological functions they provide consequently compromised or lost. Management activities are increasingly being undertaken to attempt to ameliorate this damage and to restore these functions. However, these efforts are often poorly monitored or not monitored at all and as a result we have a poor understanding of the ecological benefits that might be provided by riparian works. This lack of effective monitoring also means that there is a limited knowledge base upon which to assess the successes and failures of past projects to enhance the probability of the former in the future.

The Boorowa River Recovery (BRR) and Rivers of Carbon (RoC) projects have been initiated to improve the condition of riparian zones in the upper Murrumbidgee and Lachlan catchments, important tributaries in the Murray Darling basin. With support from landholders, the Lachlan and Murrumbidgee Catchment Management Authorities (CMAs) and other organisations, works were undertaken at a large number (~70) sites to improve the condition of riparian zones. Greening Australia has implemented a monitoring program designed by the CSIRO to assess potential changes in riparian condition in response to these works. In this report, we analyse the large data set that has been collected to date to: (1.) evaluate short-term responses to riparian works, (2.) assess the potential detectability of longer-term responses, and (3.) inform the on-going monitoring and evaluation of these projects.

### Structure of this report

This report has three main sections:

(1.) An extended analysis of BRR data, building on an earlier analyses (Hale and Reich 2012). The aim is to provide a comprehensive assessment of responses in the first six years (2008-2014) of this project.

(2.) Preliminary analyses of the RoC data to assess background variability in riparian condition, and to guide future sampling.

(3.) Examining if BRR sites could serve as proxy control sites in the RoC monitoring program.

# 3. Boorowa River Recovery (BRR) Project data analysis

The BRR project was implemented with of aim of improving biodiversity and water quality within the Boorowa River catchment through the management of native riparian management, and promoting sustainable land use more broadly.

19 project sites were established along with 19 controls. At project sites (hereafter "Treatment" sites), one of five different works types was undertaken: fencing and revegetation of erosion gullies (GFR, n = 3 pairs), structural works, fencing and revegetation of erosion gullies (GEW, n =2), fencing and revegetation of streams (CFR, n=5), willow control, fencing and revegetation of streams (W, n=4) and fencing for protection (P, n=5).

All sites in the BRR project were sampled in late spring/early summer every two years from 2008 to 2014 (Gould 2013). Sites were originally sampled using three methods:

1. At a 0.5 m cross-section within each transect to assess changes in structural vegetation.

2. Along a 100 m transect (including the riparian zone on both sides of the bank) at each site for gully/streambank erosion using the CSIRO's ephemeral stream assessment, and for structural vegetation.

3. At three locations along each transect, to assess changes in benthic macroinvertebrates (using SIGNAL scores).

## Assessing changes in structural vegetation from 0.5 m cross-sections

*Methods*

In 2012, we undertook preliminary analyses to assess potential changes in structural vegetation at the 0.5 m cross-sections at each site (Hale and Reich 2012). Initially the project was established with the aim of assessing responses at the project (treatment) sites each with a paired control site but four control sites were subsequently fenced and there were delays undertaking works at two treatment sites. For this analysis, we assigned sites to match their actual rather than intended status (i.e. by converting the four 'treated' control sites to treatment sites, and the two delayed treatment sites as controls).

We used linear mixed-effects models, as this approach offers greater flexibility and can more effectively address some of the assumptions of the commonly used alternative, repeated measures ANOVA (Logan 2010). We followed the protocols outlined in Logan (2010) to assess the potential influence of three factors: Riparian Treatment method ("Type"), Site treatment ("Treat": either treatment i.e. funded, or control) and Year, and also potential interactions between these factors. The "Treat" term provides the important test of whether Treatment and Control sites differ. The Treat*Type, Treat*Type and Treat*Type*Year terms determine how these differences change according to restoration method and through time. We fitted a series of models with different potential error terms and compared their performance using AIC scores.

Many of the variables monitored during the project were catergorical (i.e. <10%, 10-50%, 51-80%, >80%). These scores were converted into a numerical scale (i.e. 1= <10%, 2 =10-50%, 3= 51-80%, 4 = >80%) prior to analysis.

*Results*

While we detected temporal changes in almost all indicators (the “Year” term in Table 2), no significant differences were observed between Treatment and Control sites (the “Treat” term). These results are discussed in our earlier report (Hale and Reich 2012).

**Key message and recommendations**

Based on our analysis, one cross-section per site is likely to be inadequate for assessing changes in structural vegetation within the riparian zone. This is likely due to high spatial variability of structural vegetation within riparian works sites.

We recommend that a larger assessment area is used at each site, or alternatively that the number of cross-sections sampled at each site is substantially increased (6-10 cross-sections are likely to be required as a minimum).

## Assessing changes in structural vegetation from transects

*Methods*

We used the same approach i.e. linear mixed-effects models as outlined above to examine potential responses to works from transect sampling. We also explored two alternative ways to characterise potential responses: (1.) looking at relative changes between Treatment and Control sites and (2.) analysing temporal changes at Treatment sites only. The former of these is often used in monitoring programs, and essentially involves calculating the difference between Treatment and Control sites at each time period (i.e. T – C). The latter may provide a clearer picture of potential changes at Treatment sites, especially if control sites are poorly matched. Using the T-C site differences as the response variable provided comparable results to the other methods and so results are not presented here. We instead present results only from the full analysis (i.e. both treatment and control sites) and the reduced analysis (i.e. only treatment sites).

*Summary of responses to riparian works in the Boorowa River project*

Previous research has demonstrated that replanting and livestock removal is likely to result in short-term changes in groundcover (e.g. decreased bare ground and increased litter - Robertson and Rowling 2000), before the canopy begins to develop several years later (e.g. Burger *et al.* 2010; Hale *et al.* 2015). Other changes such as the development of tree hollows and accumulation of coarse wood are not likely to be observed for several decades, perhaps even longer (e.g. Vesk *et al.* 2008).

We summarise overall results from the BRR project below (Table 1 and 3, Figures 1-15). Many of the expected changes were observed. For example, bare ground was lower overall (<10-50%) at treatment sites (Figure 1). Litter was also higher at treatment than control sites, although this result was not statistically significant (Figure 5). One potential explanation for this lack of a significant result is that the scoring system is likely to only detect relatively large changes. For example, Hale *et al.* (2015) recently demonstrated an increase in litter from ~50% to ~75% cover following livestock removal and replanting, which is likely to only represent an increase from a score of 2 to a score of 3 here and thus would be difficult to detect. Similarly, we observed an overall trend (but not statistically significant) for increased groundcover at treatment sites (Figure 6).

Shrub cover was higher at treatment sites (up to 50-80% cover, Figure 7), and while this in most part likely represents replanting, results were consistent across all works types (i.e. regardless of whether they involve replanting). The probability of observing native plants flowering was also higher (>~50-75% at treatment sites - Figure 10), although there was considerable site-to-site variability both within and among the two site types.

We did not observe consistent differences in variables related to the distribution and cover of trees (e.g. Tree distribution score, Tree cover score, Shade from vegetation - Figure 11- Figure 13). This likely reflects the fact that monitoring has only been undertaken for 6 years and trees are likely not to have grown sufficiently yet to change scores. Future monitoring will be required to assess if these changes occur.

Sites where fencing was undertaken for protection purposes (P) appear to have more trees and higher canopy cover (>50%) than both their paired control sites and other works sites more generally. For example, the highest values for shade, tree distribution and tree cover were at the treatment sites in this group. These sites may also have other characteristics of higher quality riparian zones, for example, they also had the highest values for the Ephemeral Stream Assessment (Figure 14).

*Illustrating the need for control sites*

In general, our three different analysis methods (i.e. the full and reduced models, and T-C differences as response variables) yielded comparable results. However, comparing the full and reduced models provides some clear examples of the importance of including control sites in monitoring programs. For example, shrub distribution scores were higher at GEW, GFR and P treatment sites, indicating that shrub cover increased in response to works. When considering data from paired control sites (Figure 6), it appears that this is likely to be the case at GEW and GFR sites where shrub cover has remained low in the absence of works. However, shrub cover has increased at both treatment and control sites where fencing has been undertaken for protection (P), indicating that factors not related to works are likely driving increases in shrub cover at these sites (e.g. rainfall).

**Key messages and recommendations**

A number of changes have occurred following works, including:

- Decreased bare ground
- Increased litter
- Increased frequency of native plant flowering
- Increased shrub cover

Many of the longer-term changes expected following works are likely to be related to the development and subsequent growth of the canopy from replanted vegetation. As such, insufficient time has passed since works were implemented to properly assess these potential responses.

We recommend that sites are monitored again after a sufficient period (5-6 years) to assess these longer term responses. Monitoring sites in the interim is likely to be unnecessary, other than ensuring they are routinely inspected to ensure compliance (e.g. are fences intact, are plants alive) and to maintain relationships with landholders.

It is vital that control sites are included in monitoring programs, to allow potential responses to works to be evaluated in comparison with changes caused by other factors.

Table 1 Summary of results from Boorowa River Recovery riparian monitoring.

| Response variable | Difference between T and C sites? | Relevant figure | Results: full analysis | Results: T sites only |
|---|---|---|---|---|
| Bare ground score | Y | Figure 1 | • Lower in general at T sites<br>• Increases at both T and C sites in gullies | • No responses detected |
| Native grass score | N | Figure 2 | • Cover highest at gully sites (across both T and C) but decreased at both sets of gully T sites | • No responses detected |
| Pasture plants score | N | Figure 3 | • Cover highest at willow T and C sites<br>• Cover increasing at most sites through time | • No responses detected |
| Annual grass score | N | Figure 4 | • Cover higher in general in 2014 | • No responses detected |
| Weeds score | N | NA | • No responses detected | • No responses detected |
| Litter score | N | Figure 5 | • Litter higher (but not statistically significnatly) at T sites<br>• Decreases in litter cover generally across sites | • No responses detected |
| Ground cover score | N | Figure 6 | • General increase (but not statistically significant) at T sites | • Evidence of differences between years (not statistically significant) |
| Shrub distribution score | Y | Figure 7 | • Higher cover at T sites, except for P sites<br>• Strongest eeffect at gully works sites<br>• Increases at T and C sites for P works type | • No responses detected |
| Lichen score | N | NA | • No responses detected | • No responses detected |

| | | | | |
|---|---|---|---|---|
| Macrophytes | N | Figure 8 | • No responses detected | • Cover differed at T sites according to works type- lowest at GEW |
| Snags and woody debris | N | Figure 9 | • Cover highest at GFR sites in 2008, and the willow C site<br>• Decrease in cover at GFR sites between 2008 and 2014 | • Differences in cover at T sites according to works type. Highest at GFR in 2008 and P sites |
| Regen distribution score | N | NA | • No responses detected | • No responses detected |
| Native flowering | Y | Figure 10 | • Increases at T sites for CFR, GEW and GFR works. Highly variable between sites | • Differences (but not statistically significant) between works types- more likely at CFR, GEW and GFR than P or W |
| Shade from veg | N | Figure 11 | • Shade comparable between T and C sites at CFR, GFR and GEW<br>• Higher at T than C for P site, opposite for W sites (i.e. C>T) | • Differences in shade at different T sites |
| Tree distribution score | N | Figure 12 | • No overall responses but trees more abundant at GEW and P T sites than C | • Significant differences in abundance between T sites – higher at GEW and P |
| Tree hollows | N | NA | • No responses detected | • No responses detected |
| Tree cover score | N | Figure 13 | • Tree cover highest at T sites in P | • Tree cover highest at T sites in P |
| Ephemeral stream assessment | N | Figure 14 | • No responses detected | • No responses detected |
| Macroinvertebrates | N | Figure 15 | • No responses detected | • No responses detected |

# 4. Rivers of Carbon data analysis

The Rivers of Carbon project (RoC) is aimed at linking areas of native vegetation with previously rehabilitated sites to create intact riparian corridors. The overall aim is the create carbon sinks in the upper Murrumbidgee and Lachlan catchments by undertaking works to improve the extent of native riparian vegetation, and to rehabilitate waterways in the region more generally.

Sites were assessed in terms of their likely potential responsiveness to works, habitat significance, links to existing riparian vegetation, cost effectiveness, carbon sequestration opportunities and relevance to the local CMAs Catchment Action Plan. As of early 2014, 39 properties have been confirmed and a similar range of works to those in the BRR project will be implemented in the future. Some of the 39 properties have more than one location where works are being undertaken. We considered these 53 locations as independent replicates (hereafter "sites") in our analyses.

RoC sites were sampled using the Rapid Appraisal of Riparian Condition (RARC) methods developed by Jansen et al. (2005). 19 indicators are scored at four transects within each site and combined to form one overall RARC score.

We examined the RoC datasdet to: (1.) assess how do RARC scores vary between sites, (2.) assess the relative contribution of individual indicators to overall RARC scores, (3.) examine the relationship between sampling effort and precision, and (4.) evaluate what future responses may be detectable in the RoC monitoring program.

### How do RARC scores differ between different works sites?

Not surprisngly, RARC scores were considerably higher at Protection works sites (Figure 16) than the other works types. Only two sites had RARC scores >30, a wetland and a willow removal site. The four main categories – CFR, GEW, GFR and W- all had scores of ~20.

### What is the relative contribution of individual indicators to overall RARC scores?

We examined the relationship between the 19 individual indicators and overall RARC scores to assess the likely importance of different indicators. Overall, RARC scores were most strongly related to seven indicators – VegWidth, VCNativeCanopy, VCCanopy, DebrisNativeLitter, VCLayers, RARCLongConnectivity and DebrisLeafLitter (Figure 17, Table 5). In contast, three indicators- VCGroundCover, FeaturesReeds and FeaturesLargeTussockGrasses – were weakly correlated with overall RARC scores.

## Examining the relationship between sampling effort and precision

*Background*

Examining the relationship between sampling effort and sampling precision can help identify the maximum number of samples required above which precision does not significantly improve. This maximum number can be used to best allocate resources e.g. trade-off within-site sampling or sampling more sites. For the RoC dataset, this will also allow us to examine if four transects are a sufficient number of samples within sites.

*Method*

For each site, we calculated all possible combinations of 2 or more transects. There is obviously only one combination of four transects at each site, but numerous combinations of two or three transects (e.g. 1,2,3 or 1,2,4 or 1,2,5 etc). We calculated the precision for each number of transects (i.e. precision based on all combinations of two transects, then precision based on all combinations of three transects and so on). Precision was calculated by rearranging sample size Equation # 1 in Elzinga et al. (2008):

$$\text{Precision} = \sqrt{\frac{Z^2 \text{ x } s^2}{n}}$$

where:
z = the standard normal coefficient, for a 95% confidence used here, z = 1.96
s = standard deviation
n = number of samples

Low precision in these calculations represents values that are closer to the sample mean i.e. are better estimates. For comparison, a precision value of 1 when mean scores of variables are <2 represents a value within 10% of the mean. In comparison, a value of 0.8 represents a value that is closer (within 8%) of the sample mean.

*Results*

For the four indicators most strongly correlated with overall RARC scores, precision values were extremely low for 2, 3 or 4 transects (Figure 18). This suggests that current levels of replication are likely providing estimates that are very close to the sample mean and thus are adequate. This is also likely to be due to the fact that most indicators are only scored between 1 and 4. While these results may suggest that within-site replication could be decreased (e.g. to 2 or 3 transects/site), it is possible that variability will increase as vegetation develops at sites with low RARC scores initially, so we would recommend that four transects are still sampled.

## What magnitude responses may be detected in future monitoring of Rivers of Carbon sites?

*Background*

We conducted power analyses to assess the magnitude of change in RARC scores that are likely to be detectable in the future. Power analyses are a commonly used tool to help guide how many samples should be monitored. Statistical power is the probability of detecting a response, if such a response actually exists (Quinn and Keough 2002). Power analyses use information on variability among treatment units (i.e. here sites) and expected response sizes (called effect sizes) to determine the level of replication that is needed. Preliminary data can be used to provide an estimate of the variance between sites and likely magnitude of responses.

*Method*

Statistical power is proportional to the following: (Quinn and Keough 2002):

1. An estimate of the effect size- how big of a change of is interest?
2. Sample size (n)
3. Variance between sampling/experimental units
4. Significance level to be used, commonly set at 0.05.

Power analysis requires estimating an effect size and the likely variance between sampling/experimental units. Power >0.8 is typically used to indicate that a monitoring program has sufficient power that an effect will be detected.

*Setting an effect size*

While setting an effect size is a critical element of power analyses, there are few protocols for doing this in ecology (Downes *et al.* 2012). Most sites had RARC scores of ~20 before works, and ~35 at the Protected sites. We selected effect sizes ranging from an annual increase in RARC scores of between 0.25 and 2.5 over a period of 20 years.

*Setting the variance*

RARC sites have only been sampled once so no data is available to estimate temporal variance in scores. We therefore used two variances – a 5 and 10% change in RARC scores not related to any response to works.

*Analysis methods*

We followed the example outlined in Bolker (2008) to estimate the statistical power of detecting linear increases in RARC scores over the first 20 years since works, based on the effect sizes and levels of variability outlined above. At present, >40 sites are being sampled but for comparison, we evaluated power based on sampling between 20 and 50 sites.

*Results*

Our results highlight that the current monitoring program is likely to be sufficient to detect an annual increase in RARC scores of 1, even if annual variability in RARC scores unrelated to works is very high (10%, or 2 points for a site with a score of 20) (Figure 19). Over a 20 year period, this suggests that an increase from an initial score of 20 to 40 would be detectable, even with high background variability. More subtle changes (e.g. an increase from 20-30) are likely to be detectable if background variability is 5% rather than 10%.

*Some caveats with our approach*

It was necessary to make some assumptions to undertake these power analyses, and these can be updated as more information comes to hand.

First, we have assumed that RARC scores will increase linearly after works, in the absence of more detailed information about the trajectory of any response. However, this may be overly simplistic (e.g. the different degradation-recovery pathways outlined by Sarr 2002). A range of more complex, non-linear responses are possible.

Second, we have assumed that annual variance in RARC scores will stay consistent in the absence of any management. This assumption may also be overly simplistic as riparian zones are highly dynamic and there may be considerable site-to-site and year-to-year variation.

These power analyses could be re-run in the future as more data is collected to update these assumptions.

**Key message and recommendations**

- RARC scores were ~20 for the most common work types, and ~35 for sites in the Protected category
- RARC scores were strongly correlated with most sub-indicators but especially (Pearsons correlation coefficient >0.8) so with the following seven: VegWidth, VCNativeCanopy, VCCanopy, DebrisNativeLitter, VCLayers, RARCLongConnectivity and DebrisLeafLitter. Three indicators were only weakly correalted with overall RARC scores (VCGroundCover, FeaturesReeds and FeaturesLargeTussockGrasses).
- The four transects at sites provide an adequate sample size to estimate RARC scores
- Power analysis illustrates that the number of sites currently being sampled is likely to be adequate to detect even relatively small (10 points over 20 years) changes in RARC scores following works.
- Future analyses could be undertaken as more data are gathered to update assumptions that have been made in the current analysis.

## 5. Could BRR control sites be used as proxy controls in the RoC monitoring program?

*Background*

As illustrated in the BRR results section, control sites are essential so that potential responses to works can be distinguished from background variability. Control sites were not included initially in the RoC project, which means there will be issues in the future attributing temporal changes at treatment sites to works that were undertaken.

We examined the possibility for some of the BRR sites to serve as controls in future monitoring for the RoC. Each program was sampled with different methods, but in 2014 Greening Australia sampled all BRR sits within both methods. We used this data to simulate retrospective RARC scores for the BRR sites, and evaluate how well our predictions compare to observed RARC scores.

*Methods and results*

Four steps were involved in our analysis: (1.) assessing which BRR monitoring indicators best correlate with RARC scores, (2.) using these correlated indicators to predict RARC scores for the 2014 dataset (i.e. which has both BRR indicator and RARC score data), (3.) comparing our predicted to RARC scores to the actual 2014 RARC scores, and (4.) calculating retrospective RARC scores for all BRR samples.

Step 1: which BRR monitoring indicators best predict RARC scores

We examined the relationship between RARC scores and BRR indicators using multiple regression. We incorporated a range of indicators identified a priori from correlation matrices as likely to be important, and dropped these in a step-wise iterative fashion until the optimal model was found based on the Akaike Information Criteria. More than 70% of the variability in RARC scores was explained by the combination of TreeDistributionScore, ShrubDistributionScore, LitterScore, RegenDistributionScore, and MacrophytesScore (Table 6).

Step 2: using the combination of BRR indicators to predict RARC scores for 2014 BRR data

Based on Table 1, we used the following equation to predict RARC scores:

Predicted RARC = 2.5329 + (1.6323 * TreeDistributionScore) + (0.8312*ShrubDistributionScore) + (2.2076 * LitterScore) + (1.7275 * RegenDistributionScore) + (1.4883 * MacrophyteScore)

Step 3: evaluating how well predicted RARC scores correlate with actual RARC scores

Overall, predicted RARC scores were strongly correlated with actual RARC scores (Figure 20). A linear regression between the two was statistically significant ($F_{1,33} = 94.45$, $p<0.001$, $R^2 = 0.73$) and had a slope of 1 and intercept of 0, illustrating that RARC scores increased by one point as predicted RARC scores increased by 1.

To examine potential variability in this relationship, we calculated the difference between predicted and actual RARC scores for each site (Figure 21). For the vast majority of sites, our predicted RARC scores were within ±2.5 points of the actual RARC scores, indicating that our predicted values are likely to be an adequate proxy. For some sites, our predicted scores were ±5 from the actual RARC score. The likely explanation for these larger discrepancies is that some of the RARC indicators identified above (Table 5) are not being adequately represented by the set of indicators we are using to calculate predicted scores. For example, the BRR monitoring does not include a comparable variable to longitudinal continuity of riparian vegetation (RarcLongConnectivity). In particular, predicted values seem to most poorly reflect RARC scores at sites set aside for Protection (Figure 21). These sites are likely to have higher values for many of the RARC indicators not incorporated into our prediction method.

Step 4 Calculating RARC scores for all BRR samples

Our results illustrate that retrospective RARC scores could be calculated for all BRR samples using the methods outlined above. These scores are likely to provide a reasonable proxy (±2.5 points of the actual RARC scores that would have been observed). Restrospective RARC scores for all BRR sites from 2008 to 2012 were calculated and can be used to provide in conjunction with data from future monitoring to better understand likely temporal changes in RARC scores in the abscence of works.

*Potential caveats with this approach*

While these proxy controls will better enable future responses to works to be disentangled from temporal changes caused by other factors, it is important to consider some potential caveats with this approach. Ideally, control sites should be selected so that the only difference between control and works sites is the fact that works were undertaken at the later. While the RoC and BRR programs overlap in terms of the geographic area they were undertaken, there is the potential that the BRR controls could be unsuitable as controls for RoC sites. For example, there is no guarantee that these proxy sites will change in the absence of works similarly to RoC sites would have if works were not undertaken.

Also, while we detected a close match between predicted and observed sites, the match was poorer for some sites (up to a ±5 difference). To fully assess the performance of our predictions, it will be important to compare background levels of variabilty in RARC scores to the level of variability we observed between predicted and actual RARC scores. If background variability in RARC scores is <2.5 points between samples, then our predicted RARC scores are likely to be inadequate and the BRR proxies will be poor controls for the RoC program.

**Key message and recommendations**

- There is a strong relationship between RARC scores and some of the indicators monitored in the BRR project
- This relationship has been used to calculate retrospective RARC scores for all BRR sites
- In the absence of *a priori* control sites in the design of the Rivers of Carbon project BRR control sites could feasibly be incorporated as proxy controls to strengthen any inference about the response to management activities
- To fully assess the applicability of our methods, it will be important that future monitoring compares background levels of variability in RARC scores to the level of variability we observed between our predicted and the actual RARC scores. If background variability in RARC scores is >2.5 points between samples, predicted RARC scores are likely to be within the bounds of natural variability.

## References


Bolker B. M. (2008) *Ecological models and data in R*. Princeton University Press, Princeton.

Burger B., Reich P. & Cavagnaro T. R. (2010) Trajectories of change: riparian vegetation and soil conditions following livestock removal and replanting. *Austral Ecology* **35**, 980-7.

Downes B. J., Lancaster J., Hale R., Glaister A. & Bovill W. D. (2012) Plastic and unpredictable responses of stream invertebrates to leaf pack patches across sandy-bottomed streams. *Marine and Freshwater Research* **62**, 394-403.

Ezringa C., Salzer D. & Willoughby J. (2008) *Monitroing and measuring plant populations*. Bureau of Land Management, California, USA.

Gould L. (2013) Analytical case study of a large riparian rehabilitation project from an IWRM perspective - Boorowa River Recovery. In: *Masters Thesis*.

Hale R. & Reich P. (2012) Summary of monitoring results for Boorowa River recovery project. (ed U. r. f. G. Australia).

Hale R., Reich P., Johnson M., Hansen B. D., Lake P. S., Thomson J. R. & Mac Nally R. (2015) Bird responses to riparian management along degraded lowland streams. *Restoration Ecology*.

Jansen A., Robertson A., Thompson L. & Wilson A. (2005) Rapid Appraisal of Riparian Condition (Version 2). In: *River and Riparian Land Management Technical Guideline Update No. 4*.

Logan M. (2010) *Biostatistical design and analysis using R: A practical guide*.

Quinn G. P. & Keough M. J. (2002) *Experimental design and data analysis for biologists*. Cambridge University Press, Cambridge.

Robertson A. I. & Rowling R. W. (2000) Effects of livestock on riparian zone vegetation in an Australian dryland river. *Regulated Rivers: Research and Management* **16**, 527-41.

Sarr D. A. (2002) Riparian livestock exclosure research in the western United States: A critique and some recommendations. *Environmental Management* **30**, 516-26.

Vesk P. A., Nolan R., Thomson J. R., Dorrough J. W. & Nally R. M. (2008) Time lags in provision of habitat resources through revegetation. *Biological Conservation* **141**, 174-86.

**Tables and figures**

Table 2 Results of linear mixed-effects models examining responses to restoration works from the vegetation transect data. Three factors were included in these models: Riparian Treatment Type (Type: 5 levels), Site Treatment (Treat: 2 levels) and Year (Year: 3 years). Bold values indicate statistically significant results (p<0.05).

| | **Type ($F_{4,35}$)** | | **Treat ($F_{1,54}$)** | | **Year ($F_{2,54}$)** | | **Type*Treat ($F_{4,54}$)** | | **Type*Year ($F_{8,54}$)** | | **Treat*Year ($F_{2,54}$)** | | **Type*Treat*Year ($F_{8,54}$)** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Response variable** | **F** | **P** | **F** | **P** | **F** | **P** | **F** | **P** | **F** | **P** | **F** | **P** | **F** | **P** |
| Annual grasses | 2.62 | 0.05 | 1.52 | 0.22 | 49.06 | **<0.01** | 2.16 | 0.09 | 0.76 | 0.53 | 0.69 | 0.50 | 1.28 | 0.27 |
| Aquatic plants | 1.38 | 0.26 | 0.48 | 0.49 | 14.54 | **<0.01** | 0.77 | 0.55 | 1.96 | 0.06 | 1.60 | 0.21 | 0.48 | 0.86 |
| Bare ground | 0.45 | 0.77 | 1.23 | 0.27 | 3.64 | **<0.03** | 0.42 | 0.80 | 1.66 | 0.12 | 1.33 | 0.27 | 1.55 | 0.16 |
| Cryptogram | 1.28 | 0.30 | 2.67 | 0.10 | 8.98 | **<0.01** | 1.38 | 0.25 | 1.27 | 0.28 | 2.67 | 0.07 | 1.38 | 0.22 |
| Litter | 1.17 | 0.37 | 0.69 | 0.40 | 7.13 | **<0.01** | 0.38 | 0.82 | 1.44 | 0.20 | 0.68 | 0.51 | 0.70 | 0.68 |
| Native forb | 1.93 | 0.13 | 0.70 | 0.40 | 11.42 | **<0.01** | 0.71 | 0.59 | 0.88 | 0.54 | 2.05 | 0.14 | 0.99 | 0.45 |
| Native grasses | 1.74 | 0.16 | 0.19 | 0.66 | 18.10 | **<0.01** | 0.49 | 0.74 | 1.22 | 0.30 | 2.00 | 0.14 | 2.47 | **0.02** |
| Pasture plants | 11.38 | **<0.01** | 0.04 | 0.84 | 34.85 | **<0.01** | 0.52 | 0.72 | 2.86 | **0.01** | 0.45 | 0.64 | 0.69 | 0.70 |
| Shrubs | 1.07 | 0.39 | 0.01 | 0.92 | 1.50 | 0.22 | 0.51 | 0.72 | 0.51 | 0.84 | 0.68 | 0.51 | 0.95 | 0.48 |
| Weeds | 0.69 | 0.60 | 0.34 | 0.56 | 11.51 | **<0.01** | 0.14 | 0.97 | 0.74 | 0.66 | 0.23 | 0.80 | 1.21 | 0.31 |

Table 3: Results of linear mixed-effects models examining responses to restoration works from the vegetation transect data. Three factors were included in these models: Riparian Treatment Type ("Type": 5 levels), Site Treatment ("Treatment": 2 levels) and Year (4 samples). Bold indicates statistically signicant (p<0.05).

| | **Type ($F_{4,28}$)** | | **Treat ($F_{1,28}$)** | | **Year ($F_{3,83}$)** | | **Type*Treat ($F_{4,28}$)** | | **Type*Year ($F_{12,83}$)** | | **Treat*Year ($F_{3,83}$)** | | **Type*Treat*Year ($F_{12,83}$)** | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| **Response** | **F** | **P** | **F** | **P** | **F** | **P** | **F** | **P** | **F** | **P** | **F** | **P** | **F** | **P** |
| Bare ground score | 3.98 | **0.01** | 2.52 | **0.03** | 2.52 | 0.06 | 1.38 | 0.27 | 1.84 | **0.05** | 0.50 | 0.68 | 0.98 | 0.47 |
| Native grass score | 12.00 | **<0.01** | 0.02 | 0.88 | 4.28 | **<0.01** | 0.14 | 0.96 | 1.71 | 0.08 | 0.58 | 0.62 | 1.39 | 0.18 |
| Pasture plants score | 12.00 | **<0.01** | 0.03 | 0.84 | 16.21 | **<0.01** | 0.32 | 0.86 | 0.78 | 0.67 | 0.20 | 0.89 | 0.94 | 0.48 |
| Annual grass score | 1.54 | 0.22 | 0.05 | 0.83 | 31.69 | **<0.01** | 0.22 | 0.92 | 2.28 | **0.01** | 1.57 | 0.20 | 1.31 | 0.23 |
| Weeds score | 0.18 | 0.94 | 0.05 | 0.83 | 1.44 | 0.23 | 1.60 | 0.20 | 0.62 | 0.82 | 0.88 | 0.45 | 0.59 | 0.56 |
| Litter score | 1.20 | 0.33 | 2.07 | 0.16 | 15.13 | **<0.01** | 0.36 | 0.83 | 1.28 | 0.25 | 0.29 | 0.83 | 0.93 | 0.52 |
| Ground cover score | 1.18 | 0.34 | 1.32 | 0.26 | 2.33 | 0.08 | 0.29 | 0.88 | 0.35 | 0.98 | 1.44 | 0.24 | 1.17 | 0.32 |
| Lichen score | 0.15 | 0.96 | 0.15 | 0.70 | 16.20 | **<0.01** | 0.36 | 0.84 | 1.01 | 0.37 | 0.21 | 0.88 | 0.30 | 0.99 |
| Shrub distribution score* | 1.39 | 0.20 | 7.62 | **0.01** | 5.29 | **<0.01** | 0.88 | 0.48 | 1.35 | 0.20 | 0.63 | 0.60 | 1.27 | **0.25** |
| Macrophytes score** | 0.40 | 0.81 | 0.24 | 0.63 | 0.94 | 0.42 | 1.93 | 0.13 | 2.20 | **0.01** | 0.69 | 0.55 | 0.73 | 0.72 |
| Snags and woody debris | 2.83 | **0.04** | 1.10 | 0.32 | 9.27 | **<0.01** | 3.43 | **0.02** | 2.62 | **<0.01** | 0.33 | 0.80 | 1.06 | 0.40 |
| Regen distribution score | 5.88 | **<0.01** | 1.34 | 0.26 | 2.91 | **0.04** | 0.52 | 0.72 | 1.88 | **0.05** | 0.37 | 0.77 | 0.29 | 0.98 |
| Native flowering | 2.00 | 0.12 | 4.62 | **0.04** | 0.91 | 0.44 | 1.01 | 0.42 | 1.57 | 0.11 | 1.33 | 0.27 | 1.55 | 0.12 |
| Shade from vegetation | 5.55 | **<0.01** | 1.42 | 0.24 | 0.45 | 0.72 | 5.36 | **<0.01** | 0.97 | 0.48 | 0.23 | 0.87 | 1.07 | 0.40 |
| Tree distribution score | 1.00 | 0.42 | 0.88 | 0.36 | 8.04 | **<0.01** | 1.24 | 0.32 | 2.86 | **<0.01** | 0.64 | 0.59 | 1.26 | 0.26 |
| Tree hollows | 1.51 | 0.22 | 0.20 | 0.66 | 2.52 | 0.06 | 1.27 | 0.30 | 1.16 | 0.33 | 0.69 | 0.57 | 0.73 | 0.72 |
| Tree cover score | 1.96 | 0.13 | 0.24 | 0.62 | 0.62 | 0.60 | 0.84 | 0.51 | 2.05 | **0.03** | 1.01 | 0.39 | 0.49 | 0.91 |
| Ephemeral stream assessment | 3.81 | **0.01** | 0.37 | 0.54 | 3.60 | **0.02** | 0.42 | 0.79 | 1.19 | 0.30 | 0.47 | 0.70 | 0.79 | 0.65 |
| | **Type ($F_{2,19}$)** | | **Treat ($F_{1,19}$)** | | **Year ($F_{1,66}$)** | | **Type*Treat ($F_{2,19}$)** | | **Type*Year ($F_{2,66}$)** | | **Treat*Year ($F_{1,66}$)** | | **Type*Treat*Year ($F_{2,66}$)** | |
| **Response** | **F** | **P** | **F** | **P** | **F** | **P** | **F** | **P** | **F** | **P** | **F** | **P** | **F** | **P** |
| Macroinvertebrates | 2.72 | 0.09 | 0.49 | 0.49 | 31.42 | **<0.01** | 1.54 | 0.24 | 2.88 | **0.02** | 0.07 | 0.97 | 1.10 | 0.38 |

*F ratio demoninators for macrophytes = 28 and 82 **F ratio demoninators for macrophytes = 28 and 78

Table 4 Results of linear mixed-effects models examining responses to restoration works from the vegetation transect data (only from treatment sites). Two factors were included in these models: riparian works type ("Type": 5 levels), and Year (4 samples).

| | **Type ($F_{4,13}$)** | | **Year ($F_{3,39}$)** | | **Type*Year ($F_{9,39}$)** | |
|---|---|---|---|---|---|---|
| **Responses** | **F** | **P** | **F** | **P** | **F** | **P** |
| Bare ground score | 1.79 | 0.20 | 0.58 | 0.63 | 1.43 | 0.22 |
| Native grass score | 0.22 | 0.88 | 0.39 | 0.76 | 1.86 | 0.09 |
| Pasture plants score | 0.17 | 0.91 | 0.43 | 0.72 | 1.55 | 0.16 |
| Annual grass score | 0.20 | 0.89 | 1.04 | 0.39 | 1.42 | 0.22 |
| Weeds score | 3.49 | **0.05** | 1.56 | 0.21 | 1.25 | 0.29 |
| Litter score | 0.47 | 0.70 | 0.15 | 0.93 | 0.38 | 0.93 |
| Ground cover score | 0.30 | 0.83 | 2.19 | 0.10 | 1.03 | 0.43 |
| Shrub distribution score | 0.92 | 0.46 | 1.43 | 0.24 | 1.53 | 0.17 |
| Lichen score | 0.40 | 0.75 | 0.22 | 0.89 | 0.33 | 0.96 |
| Macrophytes score | 3.80 | **0.04** | 0.75 | 0.53 | 0.43 | 0.91 |
| Snags and woody debris | 4.53 | **0.02** | 0.23 | 0.86 | 1.46 | 0.20 |
| Regen distribution score | 0.70 | 0.56 | 0.36 | 0.78 | 0.37 | 0.94 |
| Native flowering | 2.38 | 0.12 | 0.92 | 0.44 | 1.44 | 0.21 |
| Shade from vegetation | 13.16 | **<0.01** | 0.30 | 0.82 | 1.16 | 0.34 |
| Tree distribution score | 4.87 | **0.02** | 0.16 | 0.93 | 1.05 | 0.42 |
| Tree hollows | 0.36 | 0.80 | 1.14 | 0.34 | 1.03 | 0.44 |
| Tree cover score | 4.02 | **0.03** | 1.19 | 0.16 | 1.55 | 0.16 |
| Ephemeral stream assessment | 0.66 | 0.59 | 0.68 | 0.57 | Not possible to estimate based on missing samples | |

Table 5 Correlation coefficients between RARC scores and 19 RARC indicators. A correlation coefficient close to 1 indicates that variables are positively linearly related, and a scatterplot of the two variables falls almost along a straight line with a positive slope.

| | RARC Score |
|---|---|
| VegWidth | 0.87 |
| VCNativeCanopy | 0.83 |
| VCCanopy | 0.82 |
| DebrisNativeLeafLitter | 0.81 |
| VCLayers | 0.80 |
| RarcLongConnectivity | 0.80 |
| DebrisLeafLitter | 0.80 |
| FeaturesNativeCanopySpeciesRegen | 0.66 |
| DebrisHollowBearingTrees | 0.65 |
| FeaturesNativeUnderstoreyRegen | 0.57 |
| RarcProximity | 0.57 |
| VCNativeUnderstorey | 0.54 |
| VCUnderstorey | 0.51 |
| DebrisStandingDeadTrees | 0.51 |
| DebrisFallenLogs | 0.44 |
| VCNativeGroundCover | 0.44 |
| VCGroundCover | 0.28 |
| FeaturesReeds | 0.23 |
| FeatuesLargeNativeTussockGrasses | 0.13 |

Table 6 Summary of optimal multiple regression model assessing the relationship between RARC scores observed at BRR sites in 2014 and five predictor variables.

| | Estimate | St Error | T value | Pr |
|---|---|---|---|---|
| Intercept | 2.5329 | | | |
| TreeDistributionScore | 1.6323 | 0.4968 | 3.286 | <0.01 |
| ShrubDistributionScore | 0.8312 | 0.4484 | 1.854 | 0.07 |
| LitterScore | 2.2076 | 0.9780 | 2.257 | 0.03 |
| RegenDistributionScore | 1.7275 | 0.6989 | 2.472 | 0.02 |
| MacrophytesScore | 1.4883 | 0.5399 | 2.757 | 0.10 |

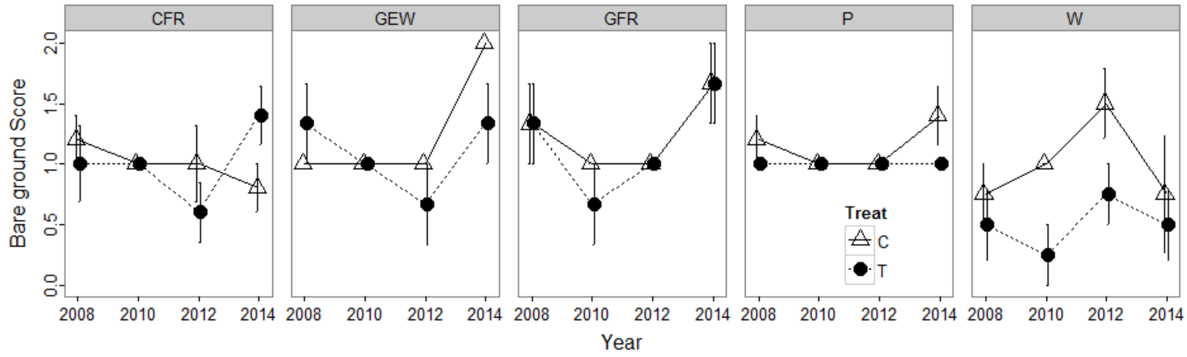


Figure 1 Changes in bare ground score at treatment (T) and control (C) sites in the Boorowa River Recovery project. Five types of works were undertaken at treatment sites: fencing and revegetation of erosion gullies (GFR), structural works, fencing and revegetation of erosion gullies (GEW), fencing and revegetation of streams (CFR), willow control, fencing and revegetation of streams (W) and fencing for protection (P) and willow control (W).

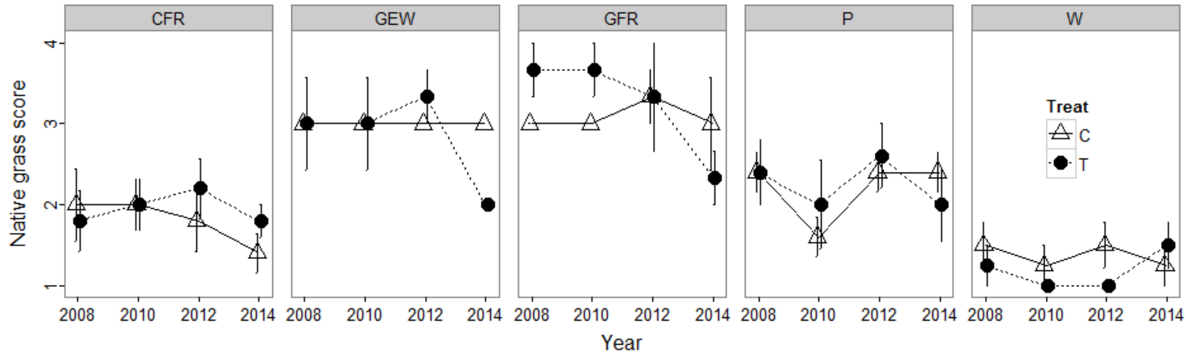


Figure 2 Changes in native grass score at treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

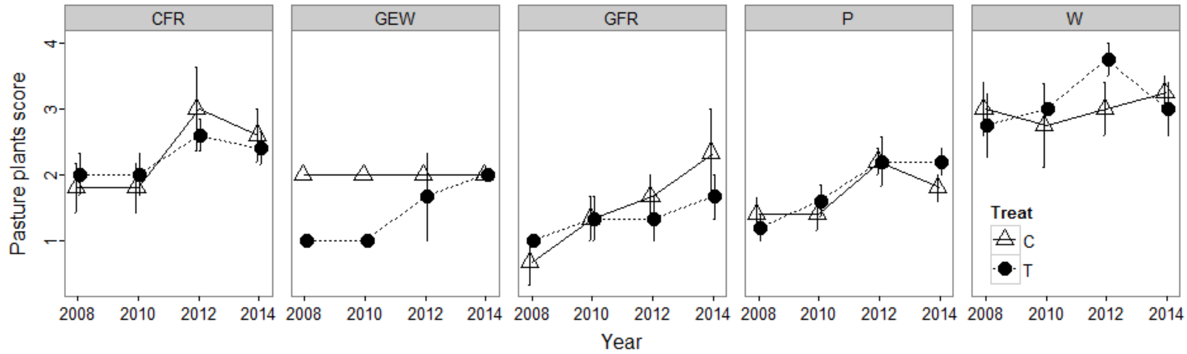


Figure 3 Changes in pasture plants score at treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

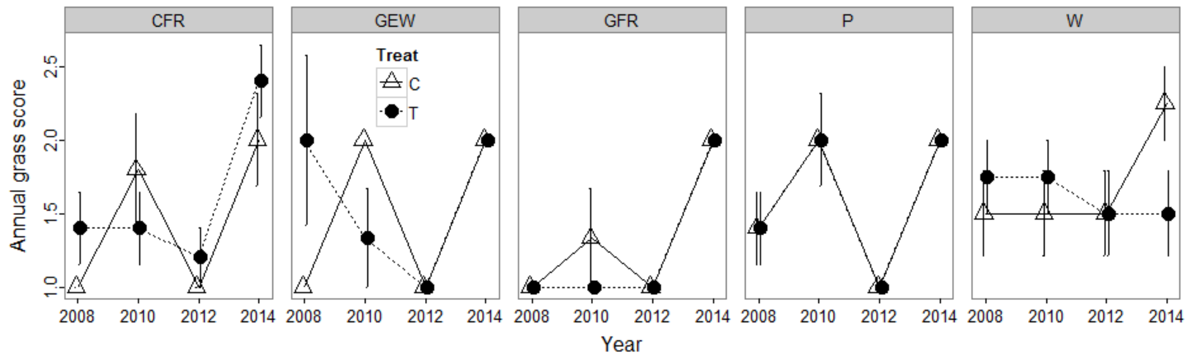


Figure 4 Changes in annual grass score at treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

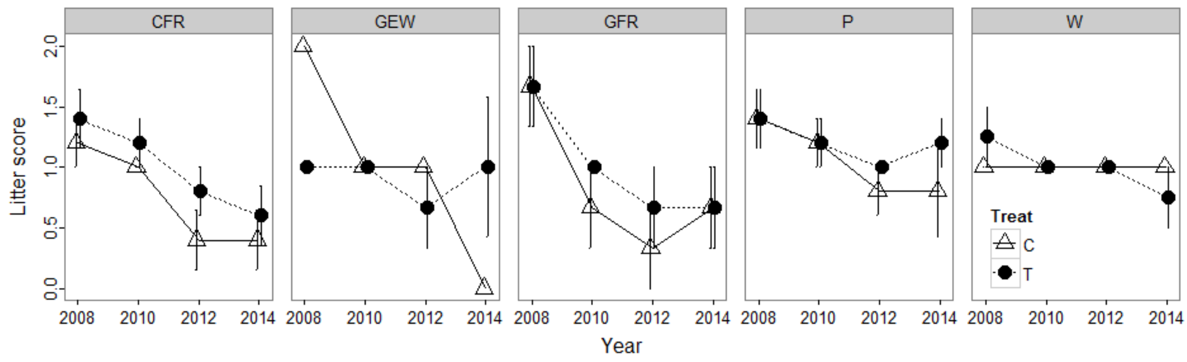


Figure 5 Changes in litter score at treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

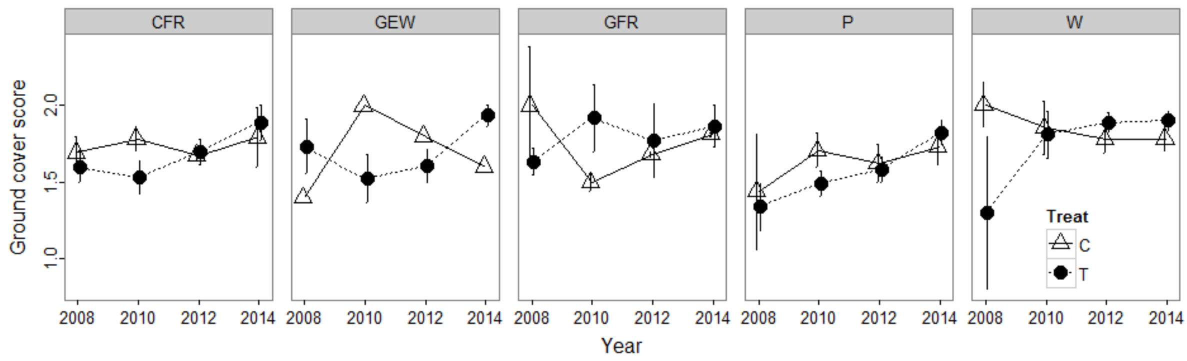


Figure 6 Changes in ground cover score at treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

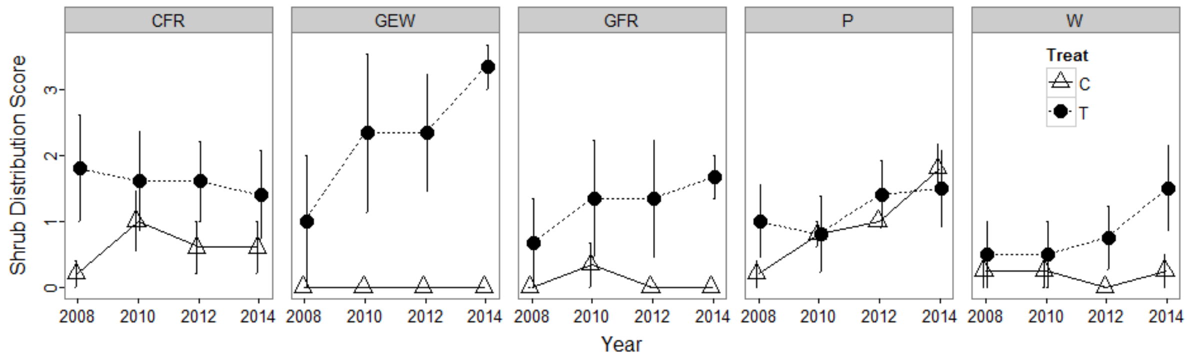


Figure 7 Changes in shrub distribution score at treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

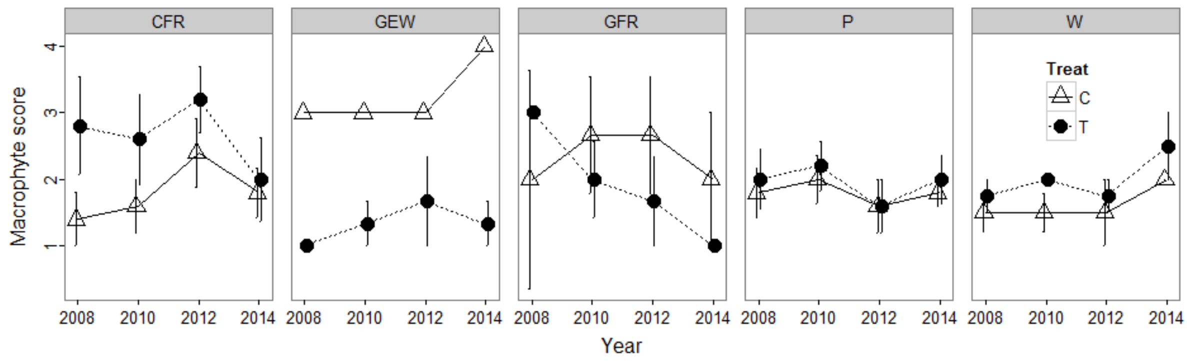


Figure 8 Changes in macrophyte score at treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

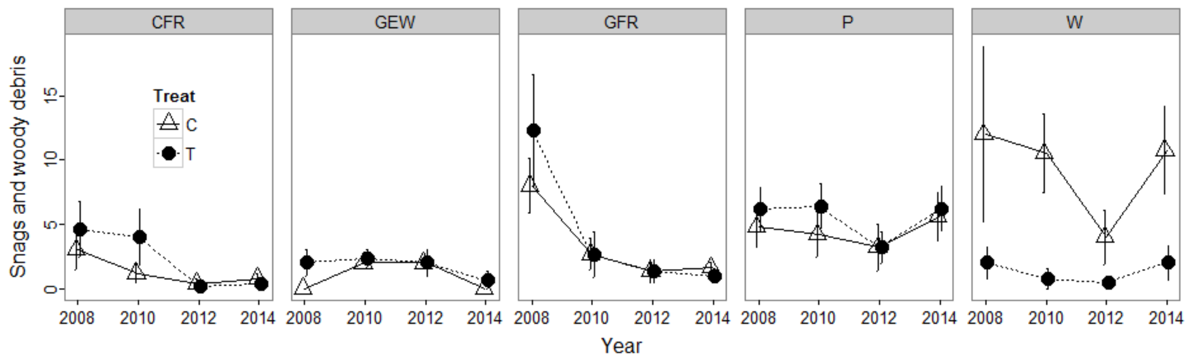


Figure 9 Changes in snags and woody debris at treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

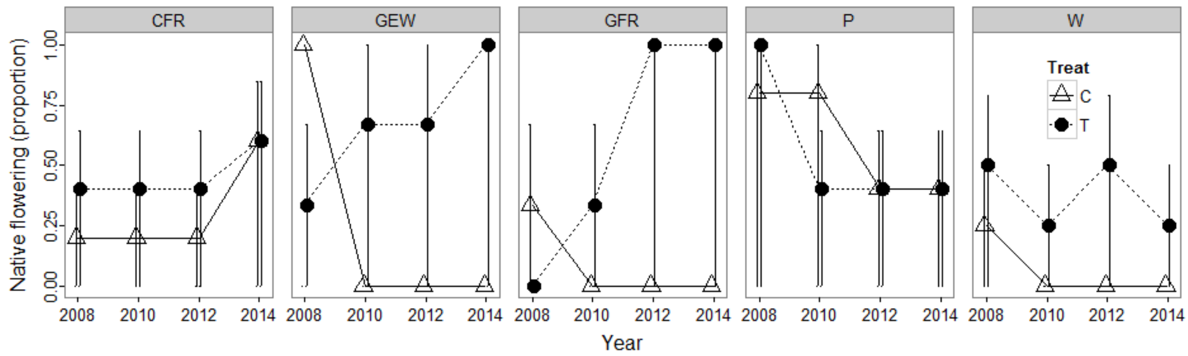


Figure 10 Changes in native flowering at treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

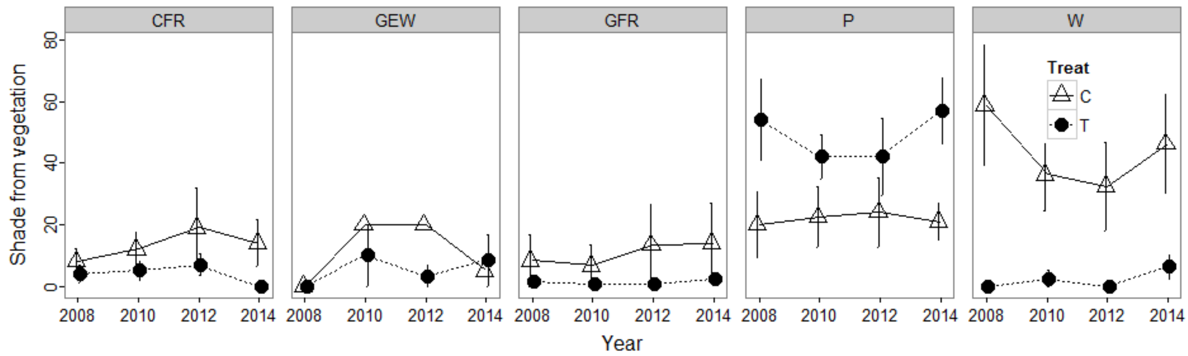


Figure 11 Changes in shade from vegetation at treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

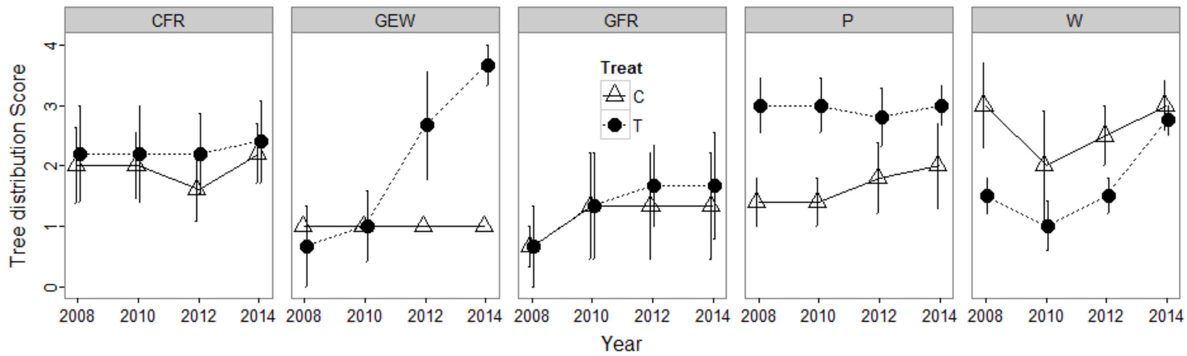


Figure 12 Changes in tree distribution score treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

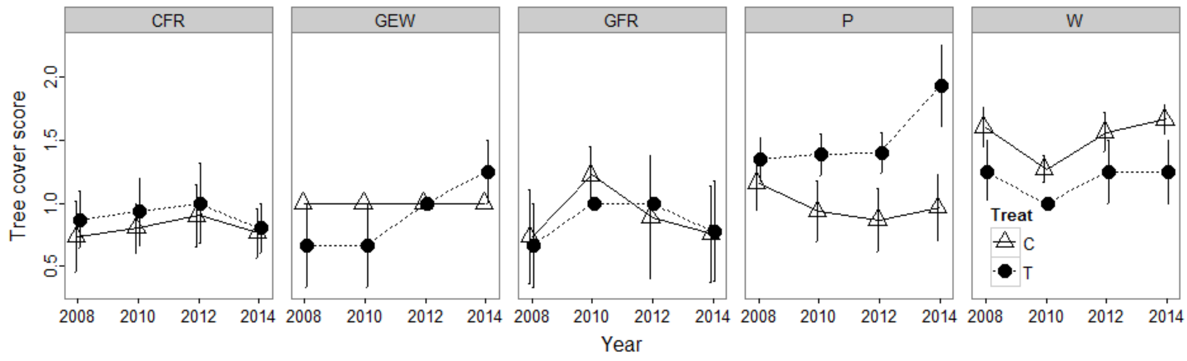


Figure 13 Changes in tree cover score at treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

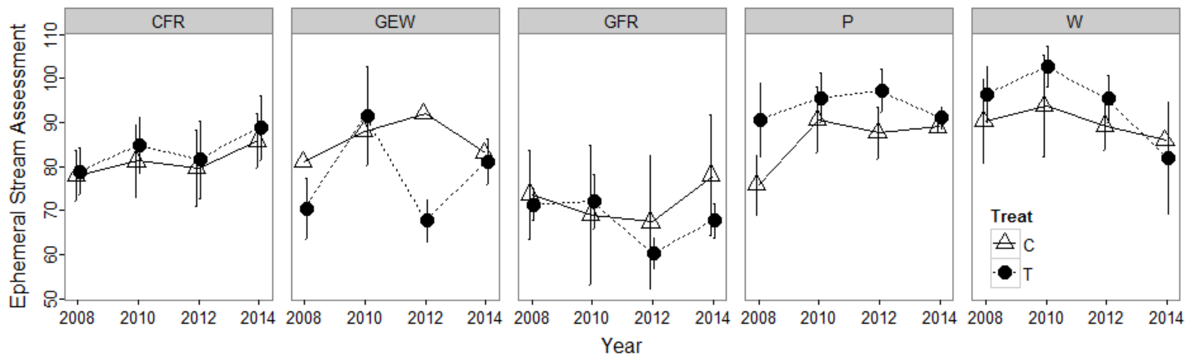


Figure 14 Changes in Ephemeral Stream Assessment values at treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

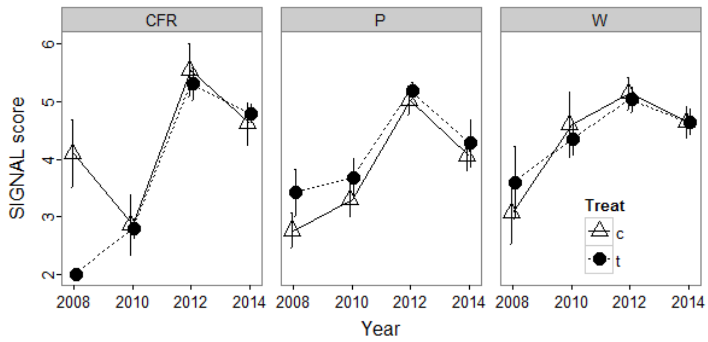


Figure 15 Changes in stream macroinvertebrates (SIGNAL score) values at treatment (T) and control (C) sites in the Boorowa River Recovery project. Figure details follow Figure 1.

Figure 16 Differences in RARC scores across the different types of works sites (mean ± standard error)

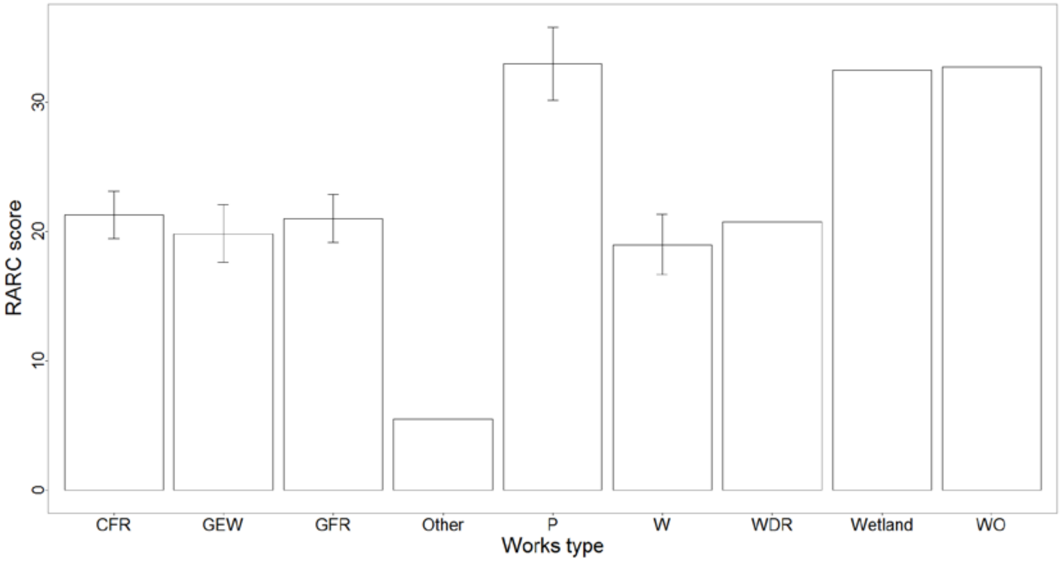

Figure 17 Scatterplots showing the relationship between Rapid Appraisal of Riparian Condition (RARC) scores and the 19 indicators used to calculate scores. Scores for the indicators were calculated as the average across the four transects monitored at each site.

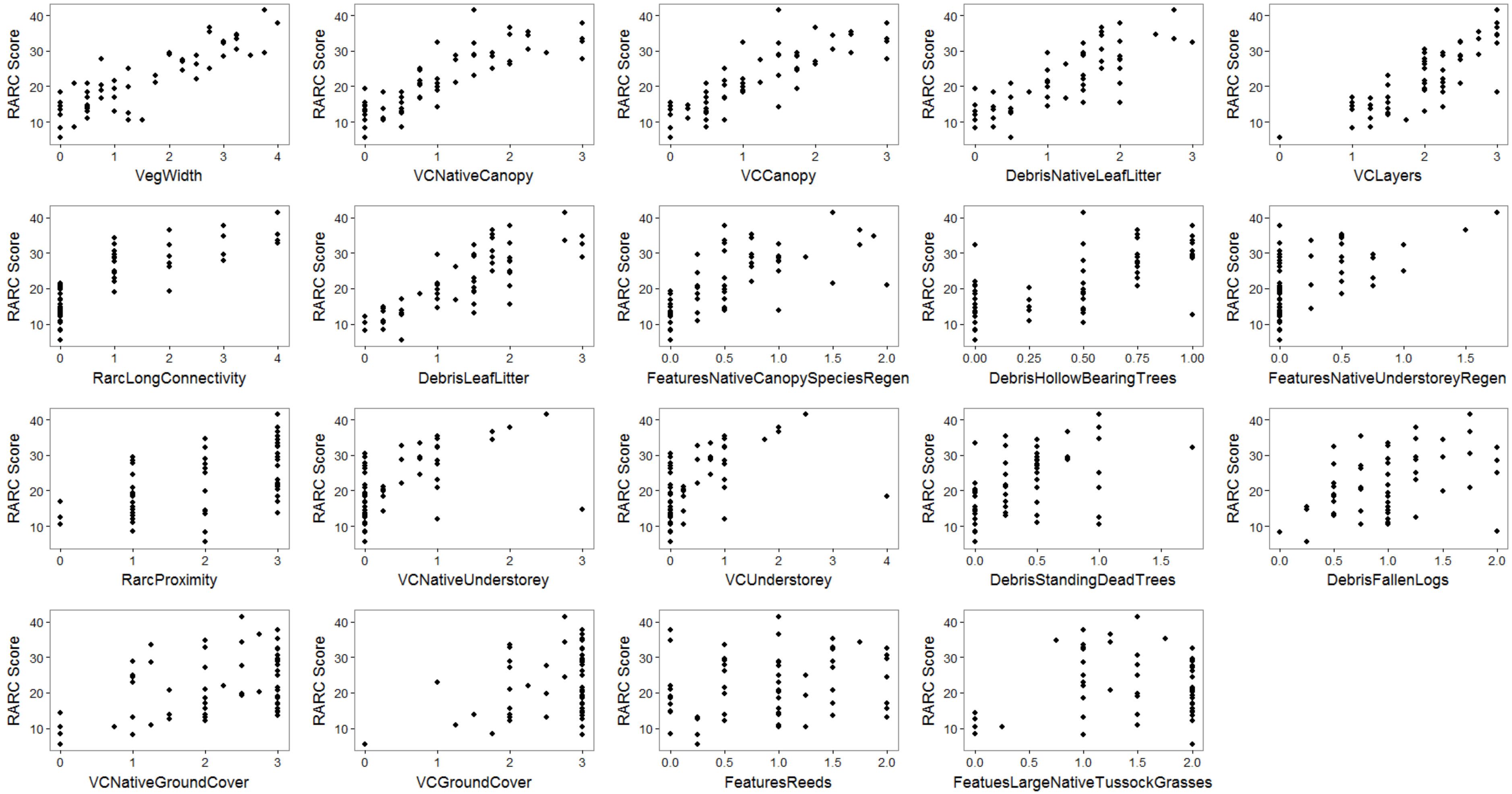

Figure 18 Relationship between precision and number of transects for (a.) VegWidth, (b.) VCNativeCanopy, (c.) DebrisNativeLeafLitter and (d.) VC Layers.

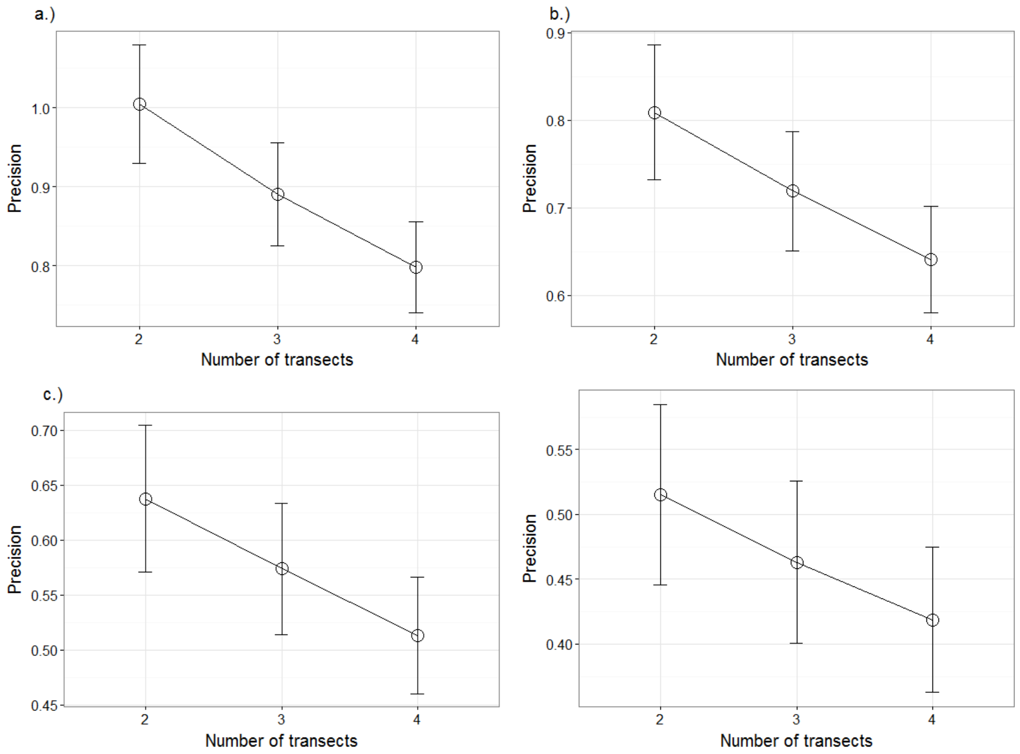


Figure 19 Power to detect changes in RARC scores in the 20 years post-works, under two levels of annual variability in RARC scores unrelated to works (a.) 5% and (b.) 10%. Effect size here refers to annual increase/decrease in RARC scores. The x-axis refers to the number of required sites.

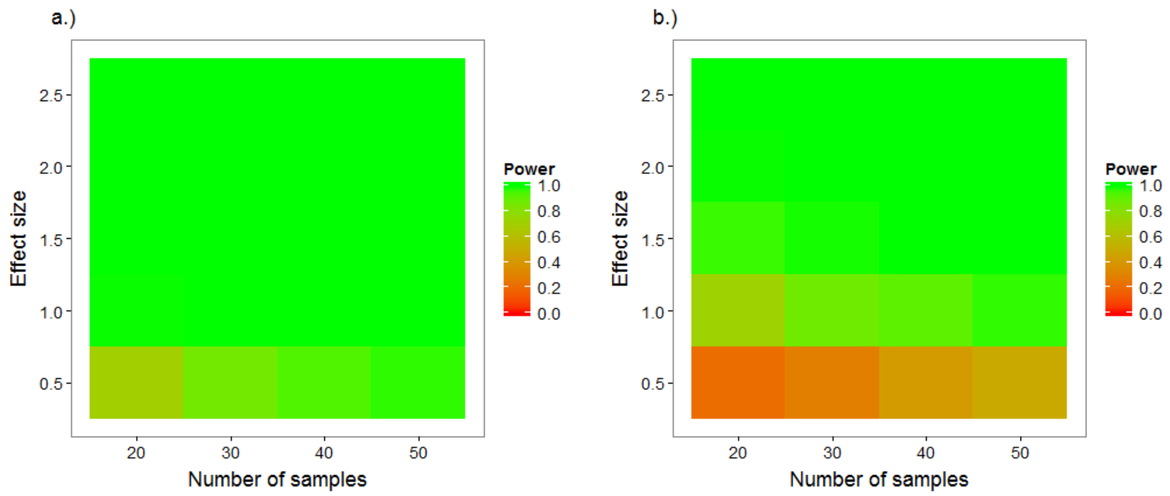

Figure 20 Relationship between actual RARC scores and RARC scores predicted from BRR monitoring indicators. The grey area highlights 95% confidence intervals based on a Loess smoothing function.

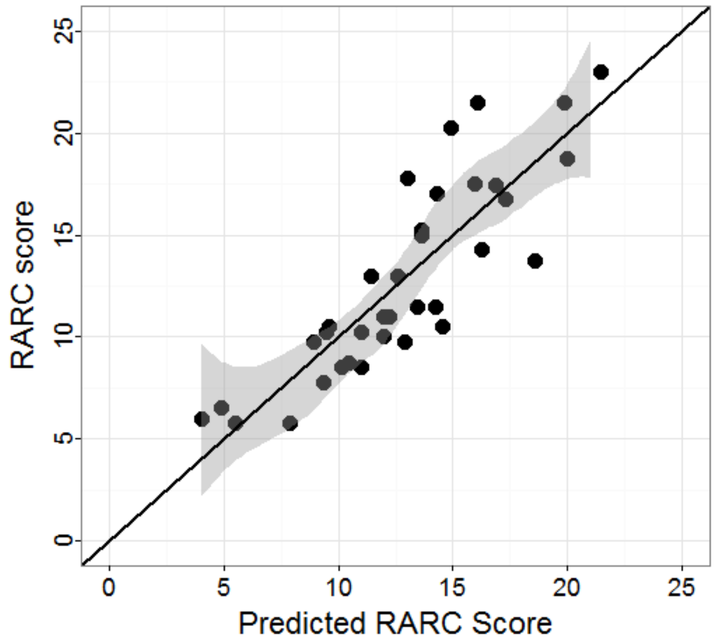


Figure 21 Examining the difference between predicted and actual RARC scores across the range of RARC scores observed in the study.