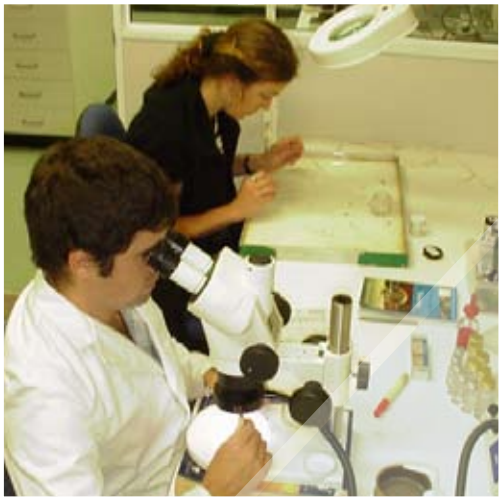
SNOWY RIVER RECOVERY

SNOWY RIVER FLOW RESPONSE MONITORING

REPRESENTATIVENESS AND EFFICIENCY OF A LABORATORY SUB-SAMPLING METHOD FOR THE SNOWY RIVER MACROINVERTEBRATE SAMPLES



SEPTEMBER 2008

NSW Government
Department of Water & Energy

Publisher

NSW Department of Water and Energy
Level 17, 227 Elizabeth Street
GPO Box 3889
Sydney NSW 2001

T 02 8281 7777 **F** 02 8281 7799

information@dwe.nsw.gov.au

www.dwe.nsw.gov.au

Snowy River Recovery
Snowy River Flow Response Monitoring:
Representativeness and efficiency of a laboratory sub-sampling method for the Snowy River macroinvertebrate samples

September 2008

ISBN 978 0 7347 5385 4

ISSN 1449 8022

Acknowledgements

Written by Matthew Russell

Andrew Brooks, Adam Boey and Simon Williams are thanked for their contribution towards the analysis and review of the report. Simon Williams edited the report.

This publication may be cited as:

Russell, M. (2008) *Representativeness and efficiency of a laboratory sub-sampling method for the Snowy River macroinvertebrate samples.* Snowy River Recovery: Snowy River Flow Response Monitoring. Department of Water and Energy, Sydney

DWE 08_198

Contents

Figures

Tables

Abstract

The macroinvertebrate component of the Snowy River Flow Response Monitoring and Modelling project requires the collection, sorting and identification of 360 samples per year. The sorting and identification of macroinvertebrates for the project is a time-consuming process and the reduction of sample processing through sub-sampling could potentially save time and money. However this could result in the reduction in the quality of collected data.

The objectives of the study were to

- to determine any differences in diversity, and community composition between two sites, in pool and riffle habitats; and
- to determine if differentiation (or not) is maintained at various sub-sampling levels (10%, 25% and 50% of the original sample); and
- to assess changes in resource/effort of sub-sampling

The results indicates that although there is an increase in variation with decreasing sub-sampling effort significant differences are still maintained at all sub-sampling levels. However it was concluded that some less abundant families can become underrepresented due to decreasing sub-sampling effort. Consistent, less abundant families can be important contributors to similarity /dissimilarity between sites and are most susceptible to underrepresentation.

It was recommended that no further macroinvertebrate sub-sampling should be conducted at this time as further sub-sampling may lead to underrepresentation of macroinvertebrate families and the assemblages of the sites of which they were collected. This in turn may affect the ability of the monitoring program to detect change to environmental flows.

1. Introduction

The Snowy River Environmental Flow Response Monitoring and Modelling Project was established to provide a physical, chemical and biological assessment of the river and quantify the changes, if any, caused by the implementation of revised environmental flow regime (EFR). The aquatic macroinvertebrate component of the program began in spring 1999 and the sampling was modified in autumn 2000 to focus on two mesohabitats (riffle, pool edges). The macroinvertebrate monitoring program aims to assess if the macroinvertebrate assemblages in the Snowy River become more similar to those in nearby unregulated rivers after the EFR compared to a regulated river with no EFR (Brooks *et al.* 2007).

On-going responsibility to provide scientific rigour to all aspects of the Project has instigated an investigation of macroinvertebrate sub-sampling. What sampling resolution is required to provide accurate, meaningful differentiation between macroinvertebrate communities from different habitats and sites? There are obvious cost implications for sampling effort, but it is also important to understand if any information loss occurs, thereby limiting the full understanding of environmental flow releases on the macroinvertebrate communities.

The sorting and identification of macroinvertebrates for the project is a time-consuming process. It takes approximately one year for one technical officer to sort and identify macroinvertebrates collected in spring and autumn each year. In an effort to reduce the time and cost a study was conducted to determine the adequacy of further sub-sampling.

This report assesses the influence of sub-sampling on the differentiation of macroinvertebrate communities between a regulated Snowy River site and an unregulated reference site.

1.1 TIME SAVINGS

The estimated time of identifying full samples compared to sub-samples are provided in Table 1. The table does not include the time to collect and sort samples as there are no savings gained from these processes. The time savings are tabulated for each habitat as pools generally have higher abundance of macroinvertebrates compared to riffles and consequently take longer to identify. Any savings should be carefully considered with respect to potential information loss from sub-sampling.

Table 1. Estimated time to identify macroinvertebrates at 10%, 25%, and 50% sub-sampling.

Habitat	Estimated time per sample	Estimated time per site	Estimated time per financial year	Total estimated time per financial year
Pool Full samples	2.5hr+	15hrs+	12.8 weeks	5.2 months
Riffle Full samples	1.5hr	9hrs	7.7 weeks	
Pools 50%	2hrs	12hrs	10.2 weeks	3.8 months
Riffles 50%	1hrs	6hrs	5.1 weeks	
Pools 25%	1.5hrs	9hrs	7.7 weeks	2.6 months
Riffles 25%	0.5hrs	3hrs	2.8 weeks	
Pools 10%	1hrs	6hrs	5.1 weeks	2 months
Riffles 10%	0.5hrs	3hrs	2.8 weeks	

1.2 OBJECTIVES

The specific aims of the study are:

- to determine any differences in diversity, and community composition between two sites, in pool and riffle habitats; and
- to determine if differentiation (or not) is maintained at various sub-sampling levels (10%, 25% and 50% of the original sample); and
- to assess changes in resource/effort of sub-sampling.

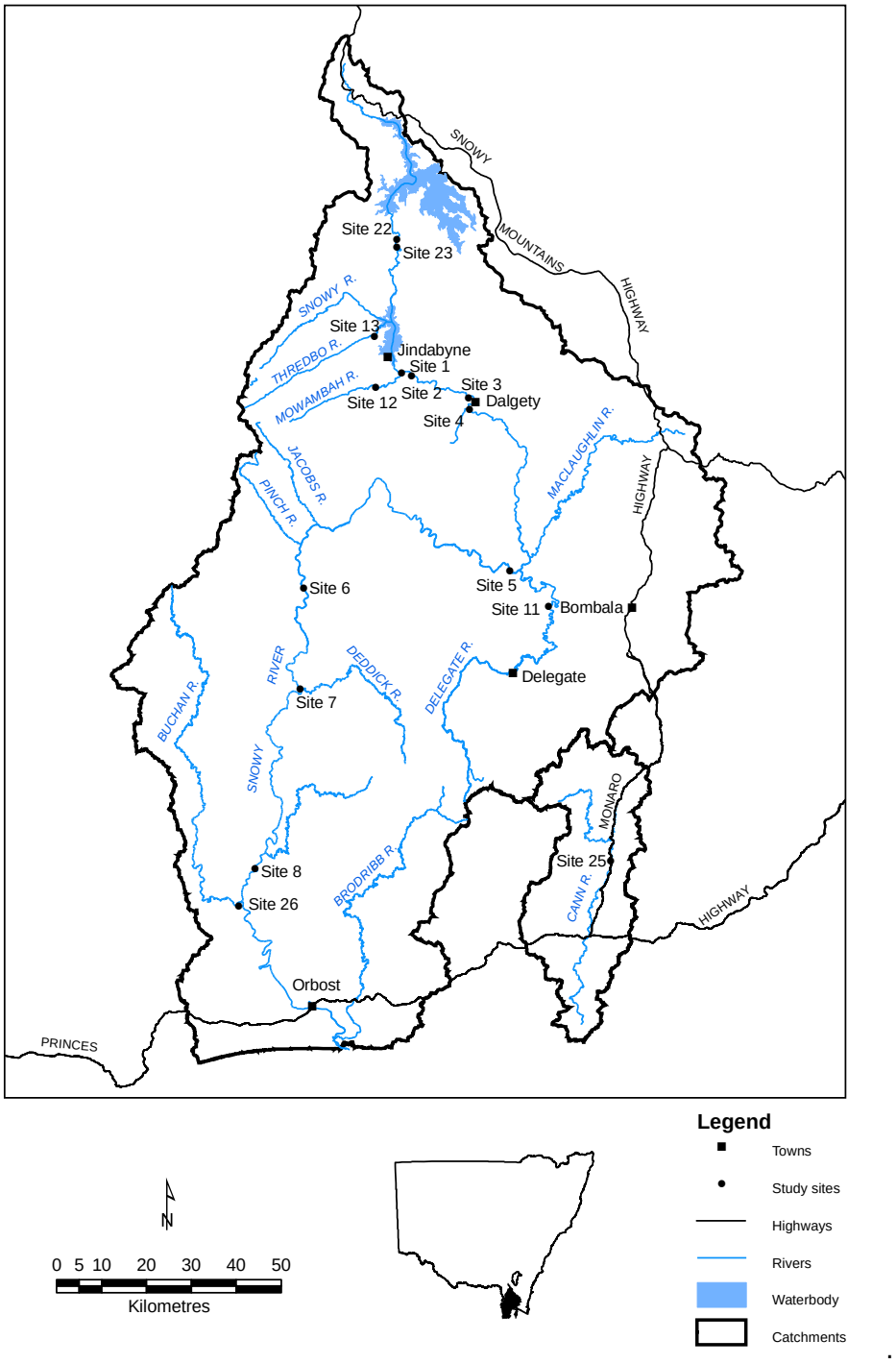
2. Study area

The Snowy River downstream of Jindabyne Dam flows for 352 km through NSW and Victoria to the river's mouth at Marlo and has a catchment area of 13,785km^2.

The Snowy River was divided into three macro-reaches for analyses based on geographic and hydrological differences. The three macro-reaches are termed upland, midland and lowland and within each one or more sites are sampled for macroinvertebrates. Each site comprises two riffle-pool sequences. Additional sites were sampled in other rivers and were used as reference and control sites for the study. Reference sites were chosen from nearby unregulated rivers and represent the ecological condition the Snowy River is expected to become more similar to with the EFR. The control sites were chosen from rivers with hydrological regimes highly altered due to regulation and will not receive environmental flows. Macroinvertebrate assemblages in the Snowy River were compared to the control sites to determine whether any biological changes observed were related to the EFR rather than region-wide influences and to also assist in quantifying the direction of faunal changes.

Snowy River at Rockwell (Site 3) and reference site Thredbo River (Site 13) were selected from the current monitoring program for this sub-sampling study. These sites were selected as they reflect a hydrological impacted and reference condition, and present a wide range of fauna. These sites will eventually be used to test the effect of the EFR.

Figure 1. Location of study sites for the macroinvertebrate sampling in the Snowy River, SE Australia.



3. Methods

Currently, six samples are collected from two pools and six samples are collected from two riffles from each site. This study aims to determine whether it is feasible to take one sub-sample from an aggregation of the six samples from each habitat/site. The methodology aims to firstly examine current differences between the impacted (by flow regulation) site, and a reference site, and secondly determine if these differences are maintained with increased sub-sampling.

The samples used were from previously collected annual autumn data from the 2000 –2004 from Snowy River at Rockwell (Site 3) and Thredbo River (Site 13). The six samples from pools were combined to form one sample and the six samples of riffles were combined to form one sample. The pools and riffles were analysed separately to allow for any differences in sub-sampling between these habitats. The macroinvertebrate data from the two sites were sub-sampled at 10%, 25% and 50% of the original samples for both pool and riffle habitats. The sub-sampling was conducted for this study using a macro program developed by Walsh (1997), which simulates the box sub-sampler described by Marchant (1989). The pool and riffle community data from the fully-processed samples were entered into the 'virtual' sub-sampler (Walsh 1997) to produce ten replicates for each sub-sampling level.

Mean plots were produced of taxon diversity for Sites 3 and 13 for pools and riffles at 10%, 25%, 50% and full sub-samples. This was used to assess difference in taxonomic representation with decreasing sub-sample proportions.

Sample by sample matrices were produced for the fully processed pool and riffle samples. This was also applied to each of the ten replicates, derived from the 'virtual' sub-sampler, at stated sub-sample levels. The similarity matrices were created by using Bray-Curtis similarity measure with abundance data transformed by log (x+1), to reduce the effect of most numerous taxa (Walsh 1997).

One-way ANOSIM was then applied to produce R-statistics for each full pool and riffle sample The R- statistic is based on rank similarities from the similarity matrix. The statistic describes the difference between the average rank similarities of all pairs of replicates between sites and the average of all rank similarities among replicates within sites.
An R- statistic of 1 implies all replicates within sites are more similar to each other than between sites while an R- statistic of 0 implies that there in no difference between sites. The statistic generally falls between these values indicating a degree of discrimination between sites (Clarke 1993).

ANOSIM was also conducted on each sub-sample replicate to produce ten R- statistics for 10%, 25%, and 50% sub-sampling proportions for both pools and riffles. The range of the ten R–statistics produced for each full sample, 10%, 25% and 50% and each median R-statistic were represented in box plots.

The Wilcoxon signed-rank test (Walsh, 1997) was used to test the null hypothesis that the median R-statistic derived for each proportion sub-sampled equalled the R- statistic of the fully processed samples. This was carried out to determine whether the differences between sites are preserved when sub-sampled. A significance level of $p<0.01$ was used to limit the possibility of Type II error (when an incorrect Ho is accepted).

Conventional measures of data quality (e.g. standard deviation) were designed for single variable estimation; insufficient or invalid for describing entire assemblages/communities (Cao *et al*. 2003). Similarity indices take species composition and relative abundance into account, and can effectively characterise community structure (Cao *et al*. 2003).

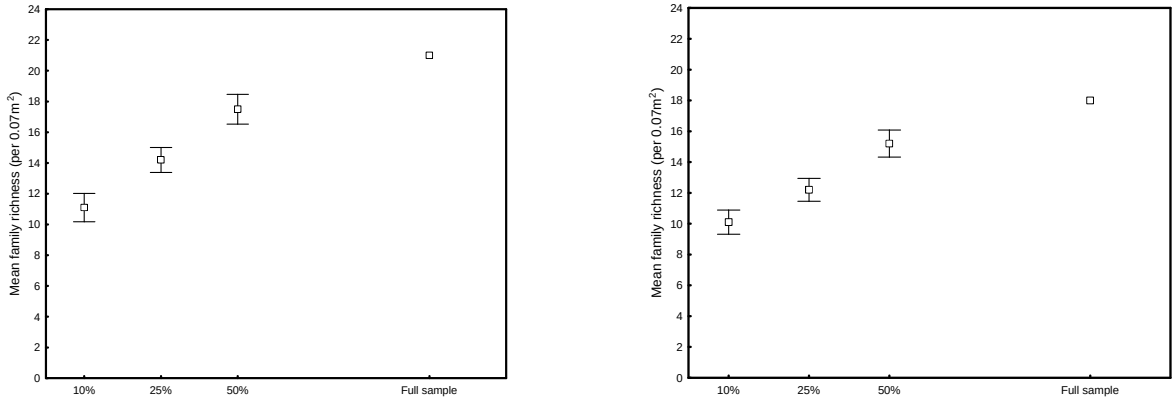
The contributions of each taxon to the overall dissimilarity between sites were quantified by the SIMPER routine (median R –statistic sub-samples for both pools and riffles to determine differences in Clarke and Warwick 1994). This was applied to the fully-processed sample, 50%, 25%, and 10% taxonomic composition.

Two dimensional ordinations of samples were produced by non-metric multi-dimensional scaling (nMDS) from the similarity matrix for the fully processed samples. An nMDS ordination was also applied to the 10% median R- statistic, 25% median R- statistic and 50% median R-statistic for pools and riffles. Where the median R-statistic lies within two sub-sample R-statistics, the lower R–statistic was used.

4. Results

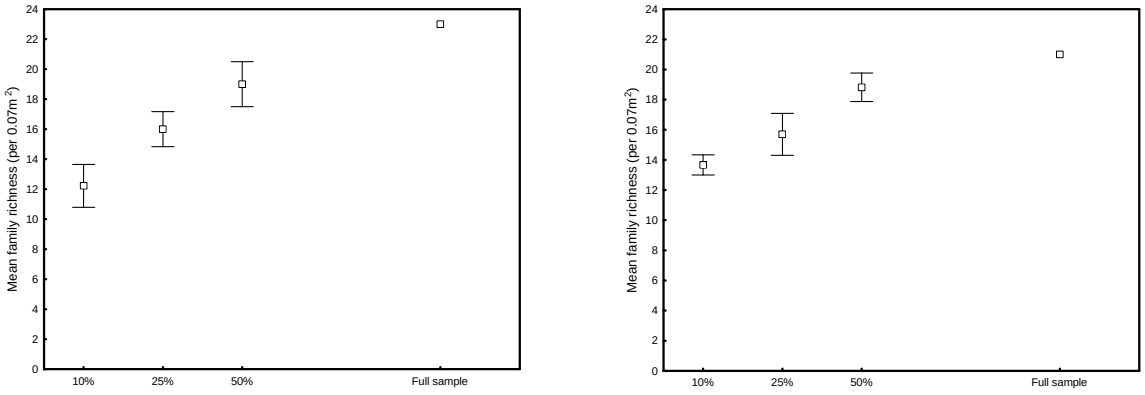
The mean diversity plots (Figure 2 and Figure 3) show considerable contrast between sub-samples and the full sample for both pool and riffles. At 10% pool sub-sampling, there is a 44% reduction in diversity for Site 3 and a 41% reduction in Site 13. At 10% sub-sampling of riffles, the reduction in diversity is 52% for Site 3 and 34% for Site 13. This signifies that less taxa are represented with decreasing sub-sample size for both habitats.

Figure 2. Pool mean family richness for the (A) Snowy River at Rockwell (site 3) and (B) Thredbo River (site13).



A **B**

Figure 3. Riffle mean family richness for the (A) Snowy River at Rockwell (site 3) and (B) Thredbo River (site13).



A **B**

The R – statistics derived from the fully processed pool and riffle samples were 0.872 (P<0.01) and 0.99 (P<0.01) consecutively, indicating significant difference between macroinvertebrate communities Site 3 and Site 13 for both habitats. The box plots in Figure 4 and Figure 5 show the range of R-statistics, of pools and riffles, for 10%, 25% and 50% proportions compared to the original samples. The ANOSIM results indicate that all sub-samples preserved community differences between Site 3 and Site 13. This is portrayed by the 10%, 25% and 50% sub-sampled median R-statistic ordinations (Figure 7).

Figure 4. Box plot of R-statistics for pool edge habitat.

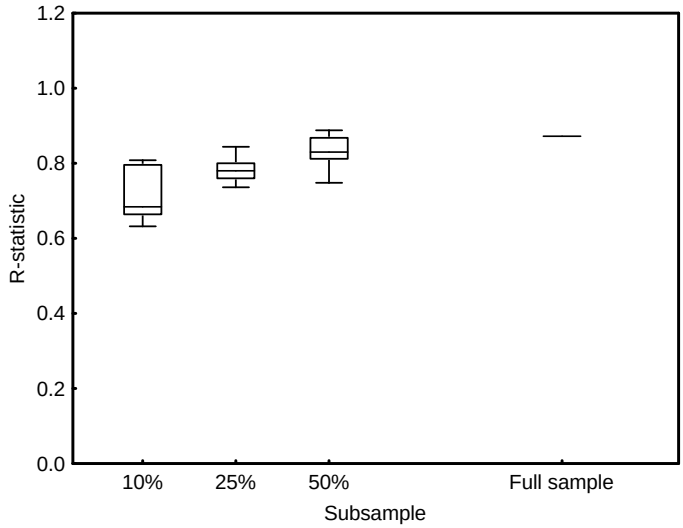
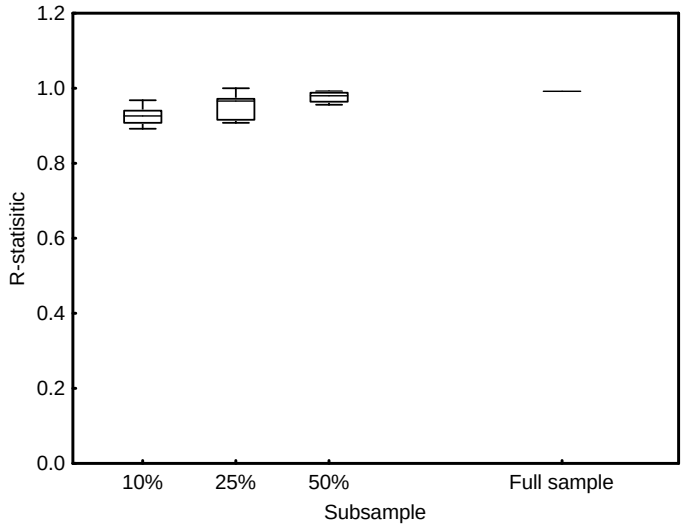


Figure 5. Box plot of R-statistics for riffle habitat.



The Wilcoxon signed-rank test indicated that there were significant differences ($p<0.01$) between the median R-statistic and the full sample R-statistic for 10% and 25% sub-samples. This occurred for both pool and riffle habitats. The 50% sub-sample median R statistic for both pools and riffles is not significantly different ($p>0.01$) (Table 2 and

Table 3).

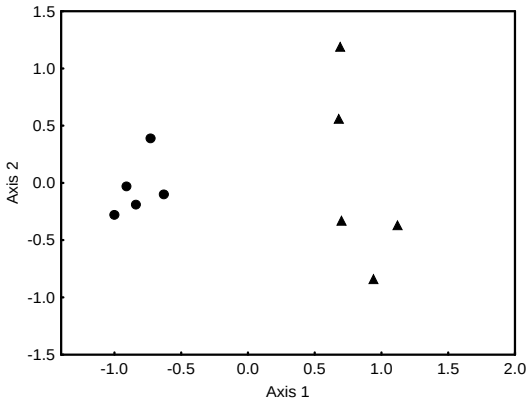
Table 2. Wilcoxon sign-rank test results for pool habitat.

Subsample	Z- Value	p-level
10%	2.803060	0.005062
25%	2.803060	0.005062
50%	2.191483	0.028418

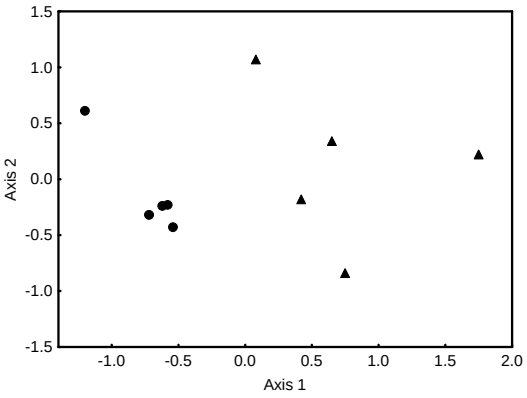
Table 3. Wilcoxon sign-rank test results for riffle habitat.

Subsample	Z- Value	p-level
10%	2.803060	0.005062
25%	2.650165	0.008046
50%	2.520504	0.011719

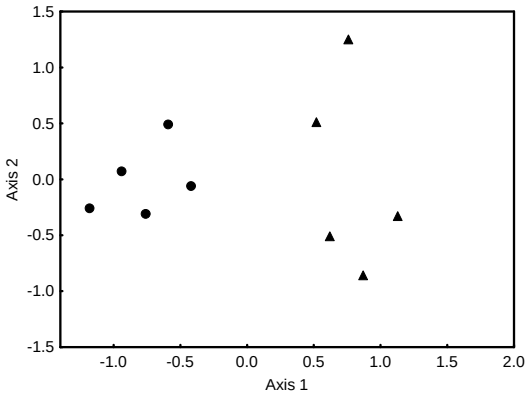
Figure 6. nMDS ordination of pool habitat for the Snowy River at Rockwell (Site 3) (●) and the Thredbo River (Site 13) (▲) for the (A) fully-processed, (B) 10% (C) , 25% (D) and 50% samples.



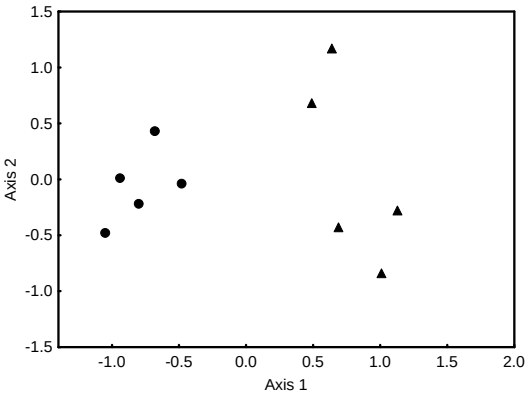
A Full sample



B 10%

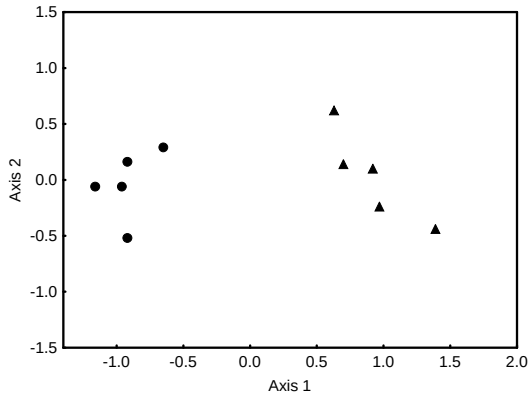


C 25%

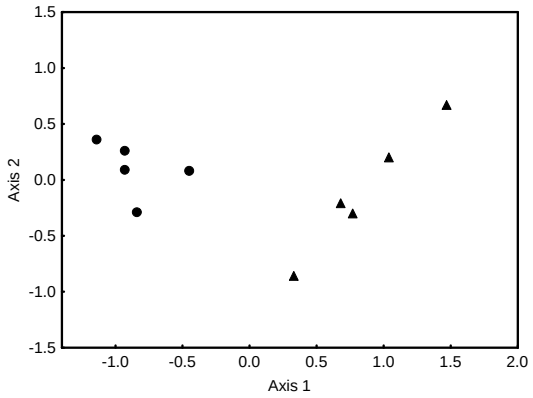


D 50%

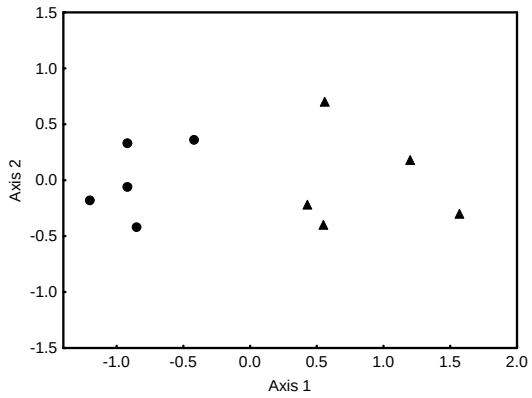
Figure 7. nMDS ordination of riffle habitat at Snowy River at Rockwell (Site 3) (●) and the Thredbo River (Site 13) (▲) for the (A) fully-processed, (B) 10% (C), 25% (D) and 50% samples.



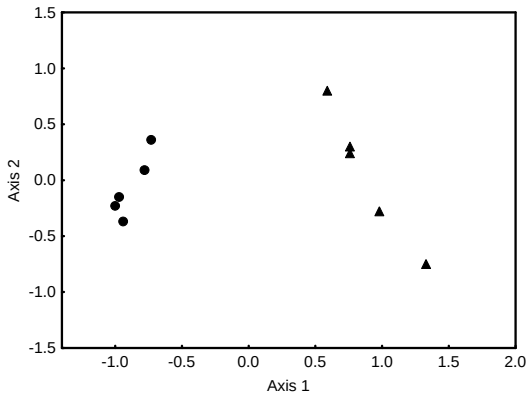
A Full sample



B 10%



C 25%



D 50%

Table 4 and Table 5 portray the contribution of taxa to the overall dissimilarity between the Snowy River at Rockwell (Site 3) and the Thredbo River (Site 13) for the full sample, 10%, 25% and 50% sub-samples of pools and riffles. It is evident that Tricladida does not contribute as strongly to the overall dissimilarity in the 10% and 25% pool samples as it does to the full sample dissimilarity. In the riffles, Pyrilidae contributes little to 10%, 25% and 50% sub-samples; however in the full sample it appears to be suitable discriminating taxa There is also a trend of less taxa contributing more to dissimilarity with decreasing sub-sampling effort for both habitats.

Table 4. Pool average abundances and dissimilarity for Site 3 and Site 13 at 10%, 25%, 50% proportions. The community data were used from the median R-statistic sub-samples and fully processed sample. * Dissimilarity/Standard Deviation.

		Site3	Site 13			
Sample	**Taxon**	**Av. Abundance**	**Av. Abundance**	**Diss/SD***	**Contrib %**	**Cum %**
10%	Oligiochaeta	165.6	12.2	3.03	10.19	10.19
	Heterodonta	29.4	6.4	1.42	7.26	17.45
	Ecnomidae	9	0.6	1.79	6.79	24.24
	Baetidae	0.2	4.6	3.53	5.76	30.01
	Caenidae	29.8	12.8	1.51	5.32	35.33
	Nematoda	8.2	0.6	1.59	5.2	40.53
25%	Ecnomidae	26.2	0.8	2.38	6.84	6.84
	Oligiochaeta	391.6	30.4	3.12	6.32	13.16
	Baetidae	0.8	12.8	2.46	5.08	18.24
	Heterodonta	63.8	15.6	1.39	4.81	23.05
	Tricladida	7.8	0	2.22	4.64	27.69
	Nematoda	14.2	0.8	2.16	4.51	32.19
	Corixidae	11.4	1	1.81	4.08	36.27
	Gomphidae	7	1	2.16	3.91	40.18
50%	Ecnomidae	50	2.6	2	5.89	5.89
	Tricladida	15.2	0	4.3	5.34	11.23
	Oligiochaeta	804.4	70.8	2.81	4.98	16.21
	Baetidae	1	22.6	3.06	4.78	20.99
	Heterodonta	125.2	28	1.43	4.39	25.38
	Corixidae	21.2	1.8	2.13	4.02	29.4
	Gripopterygidae	0.2	19.6	1.47	3.94	33.33
	Gomphidae	12.8	2.2	1.93	3.82	37.16
	Nematoda	35.6	3.4	1.79	3.46	40.62
Full	Ecnomidae	99.4	4.8	1.84	5.38	5.38
	Tricladida	29.6	0.2	5.6	5.13	10.51
	Baetidae	3	44	2.35	4.36	14.86
	Oligiochaeta	1614.8	130.6	3.27	4.16	19.03
	Gomphidae	25.8	2.8	2.23	4.08	23.1
	Gripopterygidae	0.4	37.6	1.55	3.8	26.91
	Libellulidae/Hemi/Uro**	14	0	1.75	3.54	30.44
	Heterodonta	260.8	57.2	1.39	3.5	33.94
	Coenagrionidae	12.2	0	1.74	3.47	37.41
	Elmidae	2	34.2	1.49	3.37	40.78

** Libellulidae/ Hemicorduliida/ Libellulidae are grouped as they cannot be differentiated at family level.

Table 5. Riffle average abundances and dissimilarity for Site 3 and Site 13 at 10%, 25%, 50% proportions. The community data were used from the median R-statistic sub-samples and fully processed sample. * Dissimilarity/Standard Deviation.

		Site3	Site 13			
Sample	**Taxon**	**Av Abundance**	**Av Abundance**	**Diss/SD***	**Contrib %**	**Cum %**
10%	Caenidae	138.8	0.6	3.45	12.05	12.05
	Heterodonta	45.2	0	3.71	10.57	22.62
	Ecnomidae	25.8	0.4	3.2	8.55	31.17
	Conoesucidae	0	17.2	4.04	7.62	38.78
	Elmidae	0.2	10.4	2.61	5.28	44.06
25%	Heterodonta	122.8	0	4.57	9.53	9.53
	Caenidae	360.2	3.8	2.64	9.06	18.59
	Ecnomidae	61.8	0.8	3.8	7.61	26.2
	Conoesucidae	0	40.6	5.65	7.01	33.21
	Elmidae	1.4	31.8	2.38	4.83	38.03
	Tricladida	11.8	0	1.52	4.11	42.14
50%	Heterodonta	235	0	5.13	8.97	8.97
	Caenidae	683.6	8	2.51	7.64	16.61
	Ecnomidae	128	1.2	4.43	7.07	23.67
	Conoesucidae	0.4	74.6	5.04	6.47	30.14
	Elmidae	1.8	62	3.48	4.64	34.78
	Tricladida	21.8	0.4	1.5	3.92	38.7
	Gripopterygidae	1.2	29.8	2.22	3.73	42.42
Full sample	Heterodonta	483.8	0	5.4	8.22	8.22
	Conoesucidae	0.4	156.2	6.08	6.35	14.56
	Ecnomidae	254	2.6	3.64	6.28	20.85
	Caenidae	1412.6	16.8	2.55	6.26	27.11
	Elmidae	4.2	123.8	2.77	4.02	31.14
	Tricladida	40	0.4	1.55	3.76	34.89
	Pyralidae	11.6	0	6.1	3.42	38.31
	Gomphidae	14.6	0	1.6	2.92	41.23

5. Discussion

The sub-sampling study tested whether the differences between macroinvertebrate communities of the Snowy River at Rockwell (Site 3) and Thredbo River (Site 13) were preserved when sub-sampled. The results indicated there was significant difference between Snowy at Rockwell (Site 3) and Thredbo River (Site 13) for all pool and riffle sub-samples. There was also a decrease in median R-statistic with decreasing sub-sampling effort. This indicates that with decreasing sub-sample size there is an increase in variability within sites and between sites however, community differences between the Snowy and Thredbo rivers were still significant. The reduction in R-statistic for sub-sampling that require least effort is due to increased dissimilarities between samples within treatments and are likely to be caused by differing estimates in abundance for consistent but non-abundant taxa (Walsh 1997).

The sub-sample study supports these findings as the results indicate that consistent non-abundant taxa can contribute strongly to dissimilarity between the studied impacted and reference site. These taxa also have the potential to be poorly represented due to inappropriate sub-sample sizes. These conclusions are emphasised by the reduction in diversity whereby approximately 40% of taxa are unaccounted for at 10% sub-sampling. Furthermore less taxa are contributing more to the overall dissimilarity with decreasing sub-sampling effort, indicating a change in community structure.

Further sub-sampling of macroinvertebrates may reduce the effectiveness of the analysis in detecting changes caused by environmental flows. However, this cannot be determined by this study due to the lack of post flow data as significant flows are yet to be released from Jindabyne dam. It can only be assumed that small changes to macroinvertebrate community assemblage in low abundances may not be detected due to sub-sampling and that unrepresented taxa could be important in detecting future changes caused by increase in environmental flow regime.

The implications for the Snowy River Flow Response Monitoring and Modelling Project would be to select a sub-sampling proportion that can better represent all taxa, particularly consistent non –abundant taxa by opting for a more similar sub-sample median R –statistic to the R statistic of the full sample. Walsh (1997) states that lessor sub-sampling effort result in a reduced power to detect differences, shown by the reduction in the value of R-statistics.

6. Conclusions

The following conclusions are based on the current study:

- Significant difference between Snowy River and Thredbo River samples are maintained when sub-sampled at 10%, 25% and 50% proportions.
- There is an increase in variation within samples with decreasing sub-sampling effort.
- With decreasing sub-sample size there is poorer representation of less abundant taxonomic families.
- Consistent non-abundant taxonomic families can affect similarity/dissimilarity between sites.
- Potential savings may come at the cost of the resolution of collected information.
- The magnitude of change of macroinvertebrate communities expected from flow releases is unknown and sub-sampling may reduce the ability to detect small differences from increased flows.

7. Recommendation

It is recommended that no further sub-sampling should be conducted for the Project at present, as doing so may affect the integrity of the gathered data and its ability to detect change. Further sub-sampling should be reviewed after the commencement of environmental flows to determine if sub-sampling can detect community change of aquatic macroinvertebrates under a new flow regime.

References

Brooks A, Bevitt. R, and Russell. M (2007) Response of aquatic macroinvertebrates to the first environmental flow regime in the Snowy River. New South Wales Department of Water and Energy.

Cao, Y., Hawkins, C.P., Vinson, M.R. 2003, Measuring and controlling data quality in biological assemblage surveys with special reference to stream benthic macroinvertebrates, *Freshwater Biology*, **48**, 1898-1911.

Clarke K.R. 1993, Non-parametric multivariate analyses of changes in community structure, *Australian Journal of Ecology*, **18**, 117-143.

Elliot, J.M. 1971, Some methods for the statistical analysis of samples of benthic invertebrates. Freshwater Biological Association, Scientific Publication number 25.

Marchant R**.** (1989). A subsampler for samples of benthic invertebrates *Bulletin of the Australian Society for Limnology* 12, 49–52.

Quinn G., and Keough M. 2002, *Experimental Design and Data Analysis for Biologists*, Cambridge University Press, Melbourne.

Sokal, R. and Rohlf, J. 2001, *Biometry*, W.H Freeman and Co.

Walsh, C. J. 1997, A multivariate method for determining optimal sub-sample size in the analysis of macroinvertebrate sample, *Marine Freshwater Research*, **48**, 241-48.

Warwick, R. and Clarke, K. 1993, Increased variability as a symptom of stress in marine communities, *Journal of Experimental Marine Biology and Ecology*, **172**, 215-226.