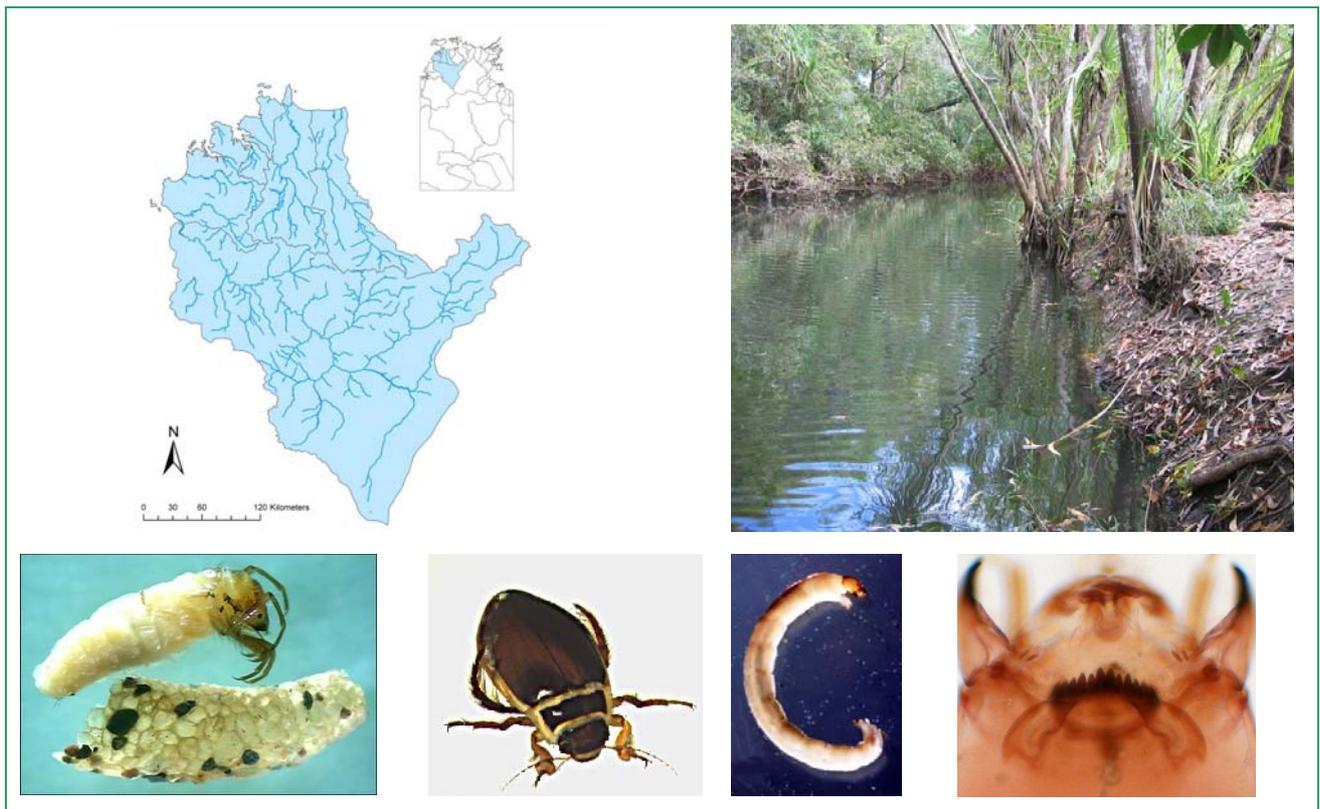# COMPARISON OF

GENUS AND FAMILY LEVEL AUSRIVAS MODELS

FOR THE DARWIN-DALY REGION

AND RELATION TO LAND USE

GISELA LAMCHE & YUSUKE FUKUDA

Report 01/2008D

Aquatic Health Unit

January 2008

Northern Territory Government

Department of Natural Resources, Environment and the Arts

Comparison of Genus and Family AUSRIVAS Models for the Darwin-Daly region and relation to Land Use.

This report should be referenced as:

CONTENTS

LIST OF FIGURES AND TABLES

1. SUMMARY

AUSRIVAS is a predictive modeling tool to assess stream health based on macroinvertebrates. This nationally standardized method is established Australia-wide, and commonly used by State and Territory governments, industry and community organisations.

A regional AUSRIVAS model was developed for the Darwin-Daly region of the NT based on mainly genus-level identification of the macroinvertebrate community.

More recently, a family-level regional AUSRIVAS model was developed for the same Darwin-Daly area that requires less taxonomic expertise. With two models to choose from, it was important to assess the sensitivity of both with respect to their ability of detecting change of the macroinvertebrate community.

The genus and the family level models for the Darwin-Daly region were compared directly based on the **o**bserved to **e**xpected ratio (O/E) scores obtained. The genus level model produced a wider distribution of O/E scores and was therefore assessed as being more sensitive. The best correlation between the genus and family level models was found on a subset of data obtained in a partly metal polluted catchment, suggesting that the family level model is as good as detecting moderate to strong impact as the genus level model.

When the two models were related to the land use in the catchment individually, the genus O/E score was explainable through horticulture and vegetation, whereas the family O/E score did not reveal a relation to any of the land use types. Overall, the genus level model allowed for some explanation of land use, whereas the family model was of limited use in this aspect.

The genus level AUSRIVAS model is hence the more sensitive of the two models and recommended if resources allow. However, there was good correlation between the two AUSRIVAS models, especially when O/E scores are low due to site impairment. The family level model is therefore also acceptable for stream health assessment.

2. INTRODUCTION

AUSRIVAS is a predictive modeling tool to assess stream health based on the macroinvertebrate community (Nichols et al. 2006, Davies 2000, Simpson & Norris 2000). The nationally standardized method is established Australia-wide and commonly used by State and Territory governments, industry and community organisations. Many regional AUSRIVAS models have been developed most of these rely on the identification of macroinvertebrates to family level to enable rapid assessment (Nichols et al. 2006, Davies 2000).

There is ongoing debate whether family level identification is sufficient to detect changes in community patterns, or whether the greater information gained from species level identification is required (e.g. Marshall et al. 2006, Arscott et al. 2006). It has been concluded that smaller scale studies, such as reaches, require species level data to discriminate patterns, whereas species assemblages over wide geographic patterns, such as several catchments, can be adequately represented at the family level (Marchant et al. 1995).

Many monitoring programs identify specimens to the resolution of family rather than species and this is also part of the rapid bioassessment methodology (Marshall et al. 2006, Simpson & Norris 2000, Wright 1995).

In the Northern Territory (NT) AUSRIVAS models based on family level were developed as part of the Monitoring River Health Initiative MRHI (Dostine 2002). These models cover most of the Top End of the NT and appear not sensitive enough to detect small impact (Dostine 2002). As many of the streams of the Top End are undisturbed or in near undisturbed condition due to the low intensity of land-use generally, more sensitive models are needed in order to detect minor deviations from the reference condition.

In 2004, a regional AUSRIVAS model was developed for the Darwin-Daly region of the Northern Territory, based on mainly genus-level identification of the macroinvertebrate community (Dostine 2004).

More recently, a second regional AUSRIVAS model was developed for the same Darwin-Daly area, but based on family-level identification in order to require less taxonomic expertise (Barlow & Lamche 2005). With two models to choose from, it was important to assess the sensitivity of both with respect to their ability to detect changes in the macroinvertebrate community.

AUSRIVAS is based on predictive models that calculate the expected compared to the observed macroinvertebrate community at a site. The results are presented in an **o**bserved to **e**xpected ratio or O/E score (Simpson & Norris 2000). The AUSRIVAS models are based on reference site data, which are defined as pristine or least disturbed sites (Davies 1994). A test site displaying a macroinvertebrate community comparable to reference sites obtains an O/E score close to one. Sites with scores higher than one are considered biologically more diverse than reference sites and are potentially biodiversity hot spots, but the high score can also be due to mild organic enrichment or continuous flow in a normally intermittent stream. O/E scores below reference sites have fewer macroinvertebrate taxa than expected, and are considered impaired. The lowest possible O/E score is zero, indicating extreme impairment of the site (Simpson & Norris 2000).

To simplify interpretation of the O/E score and to aid management decisions, a banding scheme is used representing different levels of biological condition. The interpretation of the bands is summarised in Table 1 (Coyish et al 2000).

The banding thresholds vary for each AUSRIVAS model and are shown for the Darwin-Daly regional models in Table 2.

Table 1: The AUSRIVAS banding scheme

Band	Description	O/E Taxa	O/E Taxa Interpretations
X	More biologically diverse than reference	O/E greater than 90th percentile of reference sites used to create the model.	More families found than expected. Potential biodiversity "hot-spot" or mild organic enrichment. Continuous irrigation flow in a normally intermittent stream.
A	Similar to reference	O/E within range of central 80% of reference sites used to create the model.	Expected number of families within the range found at 80% of the reference sites.
B	Significantly impaired	O/E below 10th percentile of reference sites used to create the model. Same width as band A.	Potential impact either on water and/or habitat quality resulting in a loss of families.
C	Severely impaired	O/E below band B. Same width as band A.	Many fewer families than expected. Loss of families from substantial impairment of expected biota caused by water and/or habitat quality.
D	Extremely Impaired	O/E below band C down to zero.	Few of the expected families and only the hardy, pollution tolerant families remain. Severe impairment.

The genus and the family level models for the Darwin-Daly region were compared directly based on the O/E scores obtained.

A previous study had revealed increased frequency of low O/E genus scores and land clearing in the catchment when more than 50 % of the catchment was cleared (Townsend et al. 2004). In this report, we assessed the association between the genus or family O/E scores with land clearing as well as land use parameters in the catchment.

Using the parameters land clearing or vegetation and horticulture land use in the sub-catchment, the parameters found to reveal relations to the O/E scores, another direct comparison of the genus and the family level models was carried out.

Table 2: Band thresholds for the two Darwin-Daly region models

Band	Genus level model OE50	Family level model OE50
X	> 1.14	>1.18
A	0.86-1.14	0.82-1.18
B	0.58-0.85	0.45-0.81
C	0.29-0.57	0.07-0.44
D	<0.29	<0.07

3. MATERIAL AND METHODS

3.1 AUSRIVAS sampling and data analysis

The AUSRIVAS data was collected during various projects, as well as routine monitoring in the years 1995 to 2005 by the now Aquatic Health Unit of the now Department of Natural Resources, Environment & the Arts of the Northern Territory, Australia.

Sampling and data analysis were carried out according to the user manual for the Darwin-Daly regional AUSRIVAS models (Lamche 2007).

A total of 231 O/E scores from 163 sites were calculated using each model, with some sites being sampled in several years. On a few occasions, a site was outside the respective model, leaving a total of 224 paired genus/family O/E scores.
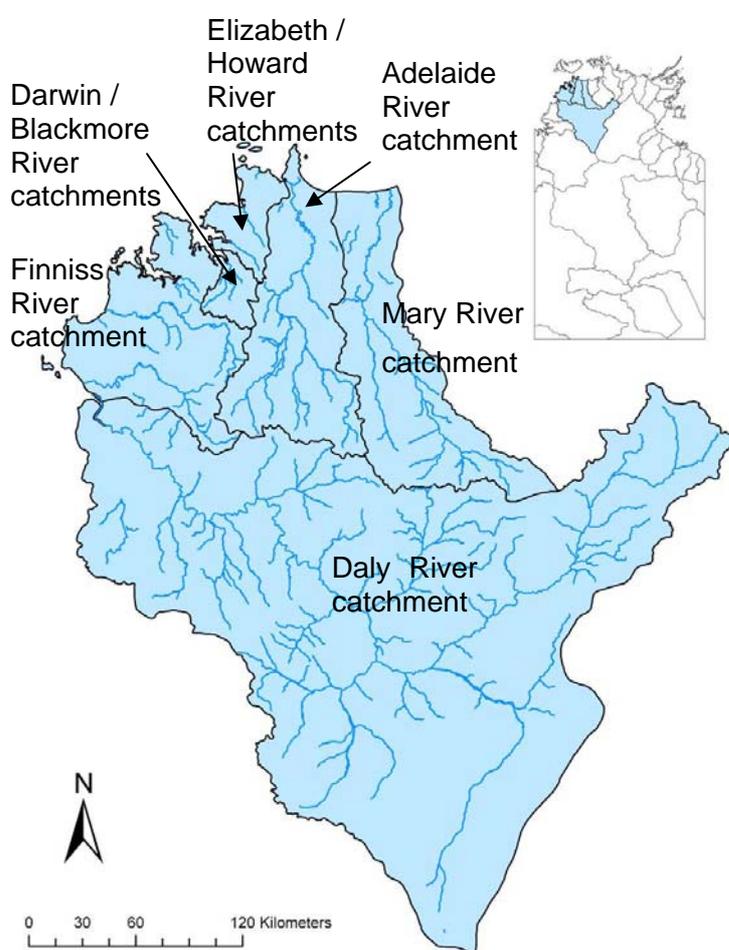
Figure 1: The Darwin – Daly region as used for the genus and family level AUSRIVAS models in the Top End of the Northern Territory.

3.2 Study areas

The area of the Darwin-Daly region as defined for the genus and family level AUSRIVAS models is shown in Figure 1 and comprises the Adelaide River, Bynoe Harbour, Daly River, Darwin Harbour (Darwin River, Blackmore River, Elisabeth River + Howard River), Finniss River, and Mary River catchments.

Data analyses of land use and O/E scores were carried out using the data for streams and rivers of the Darwin Harbour catchment as shown in Figure 2.
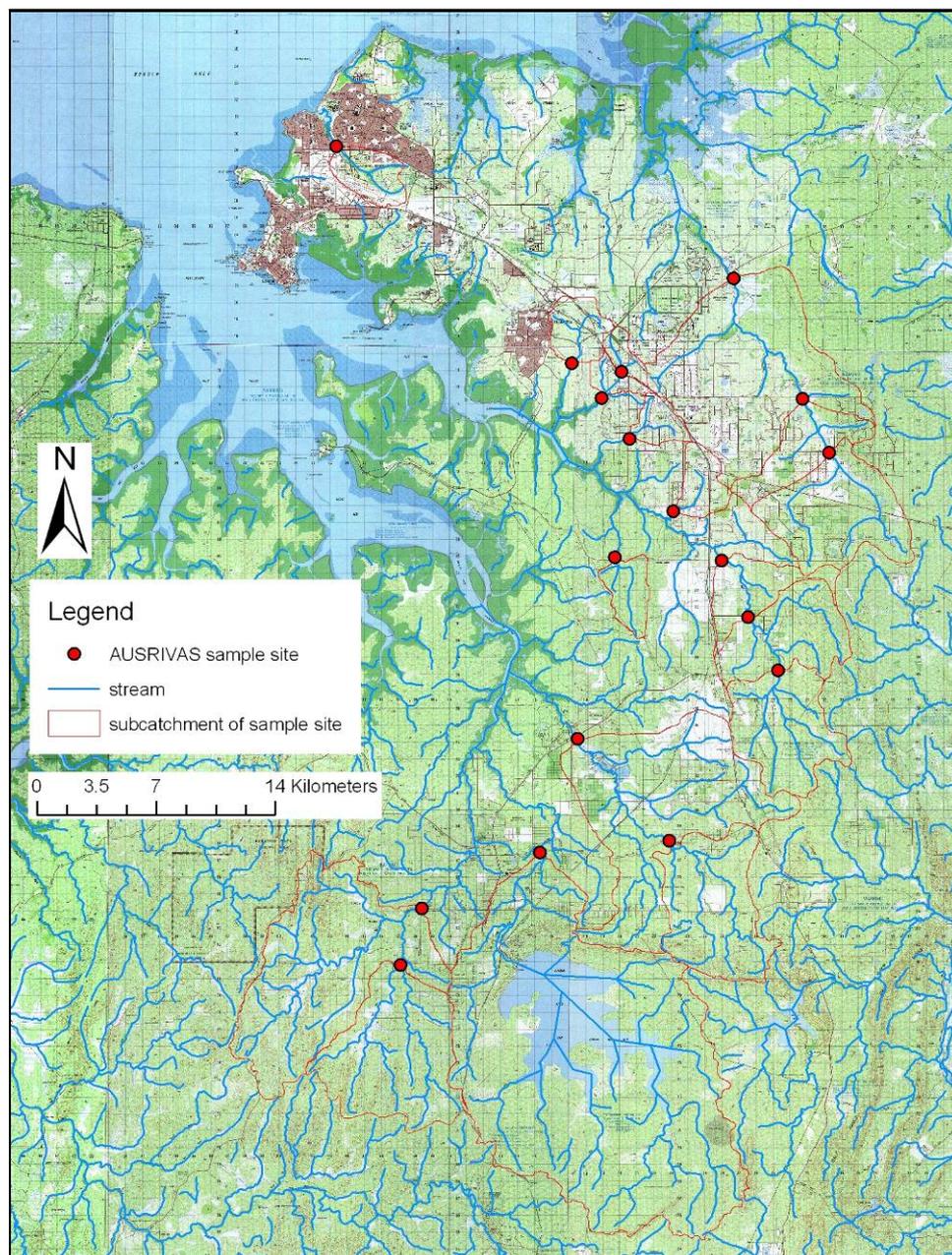
Figure 2: AUSRIVAS sample sites and sub-catchment areas in the Darwin Harbour catchment.

3.3 Land use in the sub-catchments

All spatial analysis was carried out using ArcGIS 9.1. All data sets were projected in GDA 1994.

The sub-catchment at each sampling location was obtained after plotting the sample sites. The sub-catchments were drawn based on 10 m elevation contours and 1 : 100 000 topographic data (Fukuda & Townsend 2006) (Figure 2).

Two data sets were used to obtain information in land use in the sub-catchments: 'NRETA's Land Clearing' is a vector dataset of land clearing, where vegetation has been removed. In the Northern Territory, clearing is usually the removal of native vegetation. The dataset was prepared and is regularly updated by NRETA (NT Department of Natural Resources, Environment & the Arts) based on satellite imageries at the scale of 1 : 100 000. The dataset used had been updated in 2005.

The land use map for the NT (LUMP2001) is a vector dataset on land uses in the Northern Territory. It is based on the Australian Land Use and Management (ALUM) Classification

Version 5 (Bureau of Rural Sciences 2002). The data were originally derived from a number of different sources including internal cadastral drawings, satellite imagery, aerial photograph mosaics and externally available GIS datasets. The scale of the dataset varies from 1 : 2500 to 1 : 250 000 depending on the source.

The different land use types were combined into the four following groups:

- Urban (railways, roads, rural residential, urban residential)
- Industry (manufacturing and industry)
- Horticulture (hay and silage, intensive horticulture, irrigated seasonal horticulture, irrigated sown grasses, irrigated tree fruits, perennial horticulture, sown grasses)
- Vegetation (marsh/wetland – conservation, other conserved area, other minimal use, remnant native cover).

Land clearing and land use type group areas were calculated for each sub catchment. The percentage of land clearing or land use type within the sub catchment was used as variables for the statistical modeling.

3.4 Statistical modeling

3.4.1 Candidate models

Statistical modeling was designed using the information theoretic approach (Burnham & Anderson 2001).

A dependence of the O/E scores on the land use in the sub-catchment was hypothesized. Seven models were chosen to investigate the relation of the genus level O/E score (Models 1-7) or family level O/E score (Models 8-14) with land use types.

Models 1 and 8 examine the hypothesis that the O/E value decreases with the percentage of land clearing in the sub-catchment.

The models 2 to 5 and 9 to12 use the percentage of a land use type in the sub-catchment assuming that the O/E value is lower the larger the industrial or urban land use or higher, the higher the amount of horticulture or vegetation land use.

Models 6 and 13 combine the land uses of horticulture and vegetation to investigate whether the O/E scores reveal a strong dependence on these two combined land use types.

Models 7 and 14 are a global model, incorporating all of the land use variables calculated for the various sub-catchments.

In a second approach the genus and family O/E scores were used as explanatory variables with the response variable being percentage land clearing in the sub-catchment (Models 15-16) or percentage horticulture plus vegetation land uses (Models 17-18). This allowed for the testing whether land clearing in the sub-catchment was better explained by the genus or family level O/E score (Model set 15-16). And similarly, the assessment whether the land uses horticulture plus vegetation were better explained by the genus or family level O/E score (Model set 17-18).

3.4.2 Model fitting and selection

The relationship of genus O/E and land clearing (LC) was trialed using a linear and a logarithmic fit. By plotting the response variables and visual analysis, the linear relationship was chosen to be the better fit. It was therefore decided to perform the entire modeling using linear regression.

After fitting the models, model-selection parameters were calculated for each model: Akaike's information criterion corrected AIC_c is an indication of the information that is lost when data are fitted to a conceptual model; a small (<2) delta AICc is generally considered that a model is highly supported (Burnham & Anderson 2002, Burnham & Anderson 1998); model likelihood measures the likelihood of each model within the set of candidate models; the Akaike weight is a measure of

the strength of evidence that a model is the best within the set of candidate models, ranging from 0 to 1 (0-100%).

The statistical modeling was carried out using the program 'R', version 2.1.1.

4. RESULTS

4.1 Direct comparison of O/E scores

The genus model resulted in a greater range of O/E scores from 0.36 to 1.36, the range obtained using the family level model was 0.41 to 1.30 (Figure 3). Due to the narrower range, several scores were obtained with high frequency by the family level model (Figure 3b).
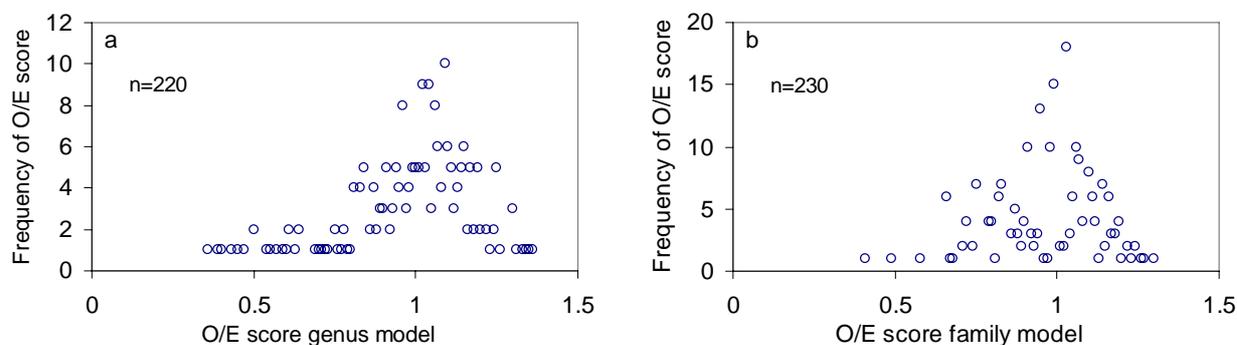
Figure 3: Frequency distribution of O/E scores obtained for genus (a) and family level (b) models for all sites.

The allocation of O/E scores into the respective Bands is summarised in Table 3. The genus level model classed fewer sites as reference (Band A) than the family model, but a considerable larger number as Band X. Both models classed 34 sites in the Band B. Only one site was allocated the Band C using the family model, whereas this Band was allocated to 10 sites by the genus level model.

Table 3: Proportional distribution of the AUSRIVAS O/E Bands for genus and family model analyses for all 231 sites.

	Genus level model	Family level model
Band A	135 (58%)	182 (79%)
Band B	34 (15%)	34 (15%)
Band C	10 (4%)	1 (0.5%)
Band X	46 (20%)	13 (6%)
Outside model	6 (3%)	1 (0.5%)

The O/E scores obtained for all sites with the genus and family AUSRIVAS models were linear correlated revealing a correlation coefficient of $R^2=0.49$ (Figure 4a). A correlation of O/E scores from reference sites only was relatively poor ($R^2=0.26$).

When choosing a subset of 24 sites within a catchment that contained reference sites and also sites impacted by point source pollution, a correlation coefficient of $R^2=0.69$ was obtained (Figure 4b).
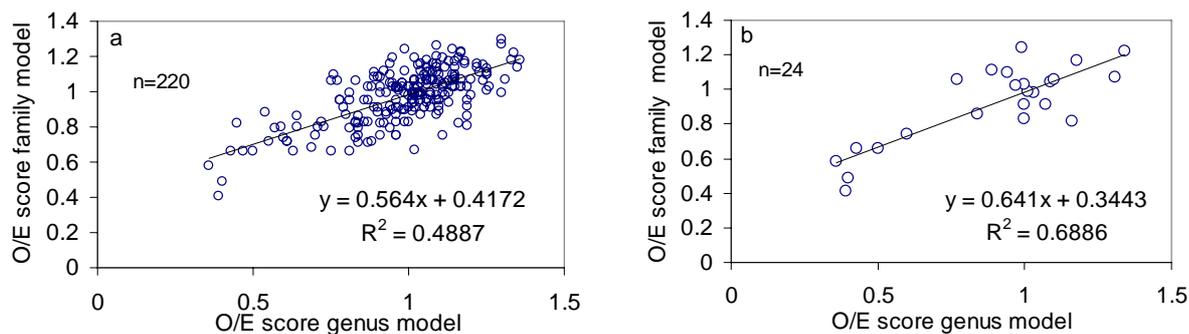
Figure 4: Plot of O/E scores obtained for family and genus level models for all 220 sites (a) and 24 sites in a catchment partly subject to heavy metal pollution (Finniss River) (b).

4.2 Calculation of environmental variables

For each of the 18 Darwin Harbour catchment sample sites the land use according to the classes above was calculated as percentage of the respective sub-catchments. The amount of land clearing in the sub-catchments in October 2005 is shown in Figure 5. The results of the spatial calculations of land use in the sub-catchments are displayed in Table 4. For the land clearing data values did sometimes vary between years. In Table 4 only the value of the last year of sampling is shown (which is mostly 2004, occasionally 2005 and for a few sites 2003 or 2002).

4.3 Environmental relationship with genus model O/E scores

All variables revealed reasonable normal distribution for all variables. These were (1) percent land clearing in sub-catchment (LC), (2) percent of manufacturing and industry in sub-catchment (industry), (3) percent of urban land use in sub-catchment (urban), (4) percent of horticulture in sub-catchment (horticulture), (5) percent of vegetation in sub-catchment (vegetation), (6) percent of horticulture and vegetation in sub-catchment (HV). Industry and horticulture did not have a normal distribution. Natural logarithmic transformation did not improve the distribution for these two parameters, nor were the other parameters distributed distinctively better. It was therefore decided to use the not transformed data.

60 genus level O/E scores were available for the 18 sample sites.

The statistical modeling was set out with the following seven *a priori* models:

Model 1: genus O/E = a * LC + b

Model 2: genus O/E = a * industry + b

Model 3: genus O/E = a * urban + b

Model 4: genus O/E = a * horticulture + b

Model 5: genus O/E = a * vegetation + b

Model 6: genus O/E = a * HV + b

Model 7: genus O/E = $(a^1 * LC + b^1) + (a^2 * industry + b^2) + (a^3 * urban + b^3) +$

$(a^4 * HV + b^4)$

with a, a^1, a^2, a^3, a^4 = estimate (coefficient); b, b^1, b^2, b^3, b^4 = intercept; LC = % land clearing in sub-catchment; HV = % of horticulture + vegetation in sub-catchment.

The parameters describing the linear fits are listed in Table 5.
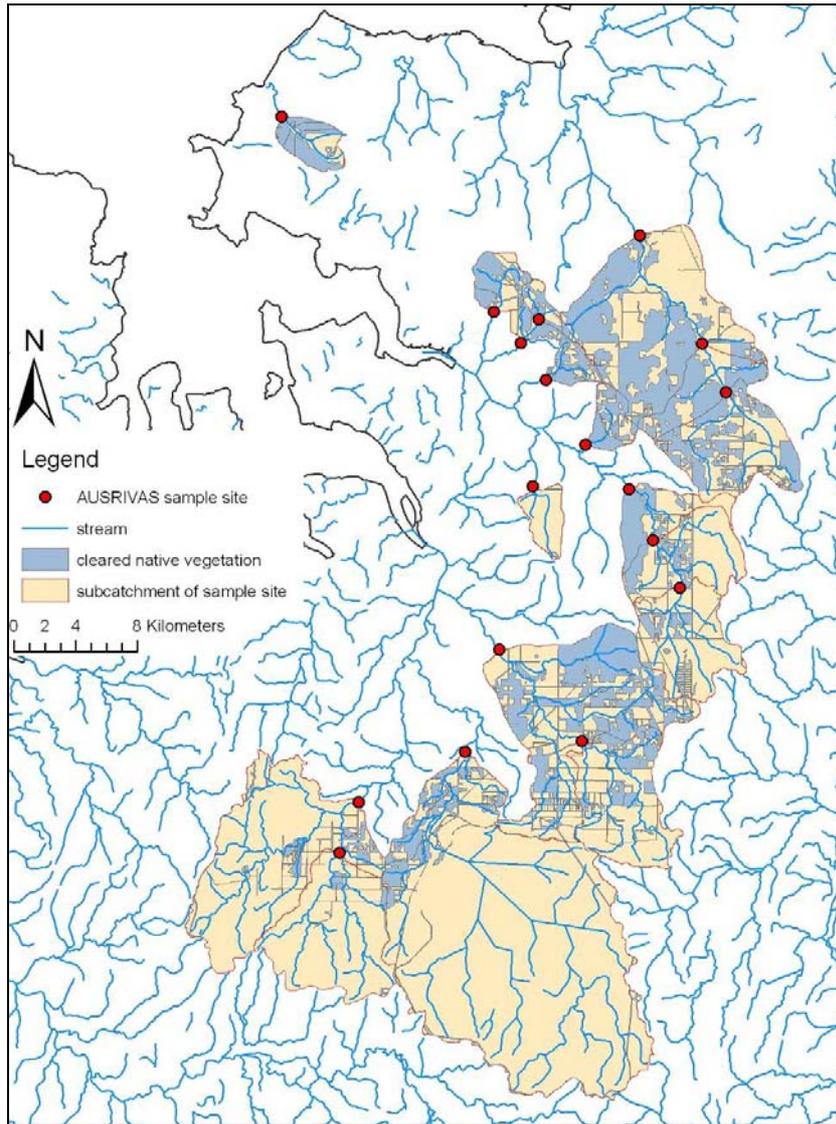
Figure 5: Areas cleared of native vegetation in each sub-catchment in October 2005.

Table 4: Various land uses expressed as percentage within the sub-catchments of each AUSRIVAS samples site.

Site number	% land clearing in sub-catchment	% urban land use in sub-catchment	% industry in sub-catchment	% horticulture in sub-catchment	% vegetation in sub-catchment
G8150127	68.66	7.53	0	0	0.88
G8155468	62.68	31.81	0.16	0	51.58
G8150102	71.30	47.99	0	11.35	9.41
G8155088	38.61	8.29	0	18.29	47.98
G8155469	5.84	13.22	0	0	83.02
G8155470	3.50	13.56	0	0	83.53
G8155471	60.19	39.70	0	3.96	41.34
G8155472	17.76	8.98	0	4.13	71.66
G8155473	24.66	8.29	0	11.70	64.59
G8155474	62.38	27.19	0	7.34	38.49
G8155475	40.07	24.95	0.10	12.90	41.45
G8150179	50.65	29.65	0.82	4.86	47.81
G8155476	16.75	8.19	0	1.80	76.19
G8155477	52.99	31.20	0.13	9.82	37.70
G8155478	22.01	4.81	0	18.20	46.35
G8155479	36.37	3.06	0	0.50	7.79
G8155510	66.77	57.41	0.64	0	15.72
G8155511	2.33	2.17	0	0	97.60

Table 5: Parameters describing the linear correlations of the 7 models explaining genus level O/E

Model number	Coefficient (a)	Intercept (b)	R^2	p-value
Model 1	1.035	-0.0031	0.16	0.00167
Model 2	0.934	-0.2794	0.12	0.00734
Model 3	0.978	-0.0036	0.08	0.02869
Model 4	0.835	0.0115	0.16	0.00188
Model 5	0.772	0.0027	0.15	0.00217
Model 6	0.713	0.0035	0.24	0.00006
Model 7			0.31	0.00038

The AICc values were calculated and listed in Table 6 with the delta AICc values and Akaike weights.

Table 6: Akaike Information Criteria corrected (AICc), delta AICcs and Akaike weights

Model number	AIC$_C$	Delta AICc	Model likelihood	Akaike weight (%)
Model 1	-37.74	6.20	0.045	2.74
Model 2	-34.93	9.01	0.011	0.67
Model 3	-32.42	11.52	0.003	0.19
Model 4	-37.52	6.42	0.040	2.45
Model 5	-37.24	6.70	0.035	2.13
Model 6	-43.94	0	1	60.75
Model 7	-42.60	1.34	0.511	31.07

4.4 Environmental relationship with family model O/E values

The checking of the distribution of the family level O/E revealed normal distribution. Natural logarithmic transformation did not improve the distribution for the family level O/E and the environmental variables. The plotting of the environmental variables versus the responsive variable family O/E did not improve possible correlations. It was therefore decided to use the not transformed data. Seven models were calculated similar to the genus level O/E modeling.

60 family level O/E scores were available for the 18 sample sites.

The statistical modeling was set out with the following seven *a priori* models:

Model 8: family O/E = a * LC + b

Model 9: family O/E = a * industry + b

Model 10: family O/E = a * urban + b

Model 11: family O/E = a * horticulture + b

Model 12: family O/E = a * vegetation + b

Model 13: family O/E = a * HV + b

Model 14: genus O/E = $(a^1 * LC + b^1) + (a^2 * industry + b^2) + (a^3 * urban + b^3) +$

$(a^4 * HV + b^4)$

with a, a^1, a^2, a^3, a^4 = estimate (coefficient); b, b^1, b^2, b^3, b^4 = intercept; LC = % land clearing in sub-catchment; HV = % of horticulture + vegetation in sub-catchment.

The parameters describing the linear fits are listed in Table 7.

Table 7: Parameters describing the linear correlations of the 7 models explaining the family level O/E values

Model number	Coefficient (a)	Intercept (b)	R^2	p-value
Model 8	1.009	-0.0026	0.1810	0.0007
Model 9	0.917	-0.1617	0.0641	0.0511
Model 10	0.971	-0.0034	0.1222	0.0062
Model 11	0.863	0.00612	0.0722	0.0380
Model 12	0.836	0.0013	0.0572	0.0657
Model 13	0.806	0.0017	0.0952	0.0165
Model 14			0.1352	0.0102

The AICc values were calculated and listed in Table 8 with the delta AICc values and Akaike weights.

Table 8: Akaike Information Criteria corrected (AICc), delta AICcs and Akaike weights

Model number	AIC_C	Delta AICc	Model likelihood	Akaike weight (%)
Model 8	-68.66	0	1	75.31
Model 9	-60.66	8.01	0.182	1.38
Model 10	-64.51	4.16	0.125	9.42
Model 11	-61.18	7.48	0.024	1.79
Model 12	-60.22	8.44	0.015	1.11
Model 13	-62.69	5.98	0.050	3.79
Model 14	-63.97	4.67	0.096	7.22

4.5 Direct comparison using genus O/E and family O/E as the explanatory variables

4.5.1 Response variable percentage land clearing (LC) in sub-catchment

The decisions above on data transformation were still valid for this modeling, therefore data were not transformed. The models tested were:

Model 15: LC = a * genus O/E + b

Model 16: LC = a * family O/E + b

with a = estimate (coefficient); b = intercept.

The combined model was included initially, but the results were not meaningful both from the modeling results as well as the interpretation point of view.

The results of the direct comparison of these two models are shown in Tables 9 and 10.

Table 9: Parameters describing the linear correlations of the two models

Model number	Coefficient (a)	Intercept (b)	R^2	p-value
Model 15	87.21	-50.57	0.1579	0.00167
Model 16	103.63	-69.08	0.1810	0.0007

Table 10: Akaike Information Criteria corrected (AICc), delta AICc and Akaike weights

Model number	AIC_C	Delta AICc	Model likelihood	Akaike weight (%)
Model 15	543.80	1.67	0.435	30.31
Model 16	542.13	0	1	69.69

4.5.2 Response variable of horticulture and vegetation as percent in sub-catchment

The decisions above on data transformation were still valid for this modeling, therefore data were not transformed. The models tested were:

Model 17:   VH = a * genus O/E + b

Model 18:   VH = a * family O/E + b

with a = estimate (coefficient); b = intercept.

The results of the direct comparison of these two models are shown in Tables 11 and 12.

Table 11: Parameters describing the linear correlations of the two models

Model number	Coefficient (a)	Intercept (b)	R^2	p-value
Model 17	-7.36	68.45	0.24	0.00006
Model 18	5.145	54.94	0.095	0.0165

Table 12: Akaike Information Criteria corrected (AICc), delta AICcs and Akaike weights

Model number	AIC_C	Delta AICc	Model likelihood	Akaike weight (%)
Model 17	548.68	0	1	99.48
Model 18	559.19	10.51	0.0052	0.52

5. DISCUSSION

The wider distribution of O/E scores for the same sites suggests the genus level model to be of greater sensitivity than the family level model, as there is more differentiation between sites. It is assumed that the genus level data better represents the macroinvertebrate community than the same data aggregated at the family level, as information is lost during the aggregation process. In general, finer taxonomic resolution is considered to have greater information content and results in more reliable site classifications (Marshall et al. 2006, Arscott et al. 2006, Waite et al. 2004). This was especially found for communities with a high proportion of chironomids (Arscott et al. 2006, King & Richardson 2002, Waite et al. 2004), and is also found for the NT community data used in this study. Small-scale studies, such as reaches, require species level data to discriminate patterns, whereas species assemblages over wide geographic patterns, such as several catchments, can be adequately represented at the family level in temperate Australia (Marchant et al. 1995).

Arscott et al. (2006) studied the macroinvertebrate communities of two stream systems with different disturbance gradients. In the catchment with the strongest disturbance gradients, family level identifications were adequate to separate the sites as were species and genus level identifications. In the catchment with the subtle gradients genus and species identifications were necessary to differentiate among sites.

Arscott et al. (2006) concluded that for a study with the primary goal to distinguish severely degraded from healthy sites and when financial resources are limited, family-level taxonomic resolution may be appropriate.

The O/E scores obtained using both of the AUSRIVAS models were correlated. In an ideal world, the same data set used for the building of both models, each of the sites should receive similar scores, independent of the model used. For example an undisturbed site would be expected to obtain an O/E score around 1 regardless whether the genus or the family level model was used for data analysis. Reasonable well correlation between the O/E scores of the two models was obtained.

Better correlations were obtained, when a subset of sites displaying a wide range of O/E scores was used such as presented in Figure 4b.

In an AUSRIVAS study in the outer urban Melbourne area, species and family level models were used (Breen et al. 1999) and with a correlation between the two models of $R^2=0.80$. In this study the range of scores was also very wide, confirming that this produces better correlations when using models of the same dataset on two taxonomic levels. The study also demonstrated that two models based on different taxonomic level even on the same dataset do not produce identical scores, in agreement with the results of the present study.

Nichols et al. (2006) demonstrated a higher variability in O/E scores in streams that were significantly impaired (Bands B to D). It was assumed that the remaining habitats in the impaired streams could be patchy, and the depauperate invertebrate fauna was distributed more thinly (Nichols et al. 2006).

In the present study, the genus level model is more sensitive by separating the sites more widely. For the subset of impacted sites, family and genus level models were better correlated, indicating that for studies on moderately to severely impacted sites the family model is as suitable as the genus level model. This subset of impacted sites was in the Finniss River catchment, which was subject to point source metal pollution (Dostine 2002). The genus as well as the family level model both classed sites according to the downstream gradient. This agrees with a study by Sloane & Norris (2003), where the AUSRIVAS family level model used explained well the metal pollution gradient of the Molonglo River. In our study, however, the genus level model did distinguish between the impacted sites better than the family level model, providing evidence that the genus level model is more sensitive to detect impact.

5.1 Relationship between genus model O/E values and environmental parameters

There is a sixty percent strength of evidence that model 6, which is a linear correlation between the genus O/E values and horticulture + vegetation in the sub-catchment describes the data best within the set of models tested. Model 6 had a correlation coefficient of 0.24, which is slightly below accepted low values indicating a correlation between variables, but is highly supported statistically. Model 7, the global model, displays a thirty percent strength of evidence within the model set. It had a correlation coefficient of 0.31 and is statistically also supported.

This can be interpreted as the land use type horticulture + vegetation has the most profound effect on stream health assessed with the Darwin-Daly genus-level model.

5.2 Relationship between family model O/E values and environmental parameters

There is seventy percent strength of evidence that Model 8, a linear regression between the family O/E and percent of land clearing in the sub-catchment, represents the data best within the model set used. However, the correlation coefficient was 0.18 and although highly statistically supported this does not allow for a correlation of the two parameters to be assumed. Thus, the family level O/E values can not be well explained by any of the seven models tested.

Family level O/E scores can not be explained by land use in the catchment.

5.3 Direct comparison using genus O/E and family O/E as the explanatory variables

5.3.1 Response variable as percentage of land clearing in sub-catchment

The direct comparison of the genus and family model O/E scores was modeled using the percentage of land clearing in the sub-catchment as the response variable, as a previous study had revealed a weak association between the genus O/E score and land clearing in the catchment above a threshold value (Townsend et al. 2004).

The family O/E variable is the better variable in explaining the responsive variable land clearing in sub-catchment, with about seventy percent strength of evidence. The genus O/E model is still contributing thirty percent. However, the fit of both linear models is well below good with a correlation coefficient of 0.16 and 0.18, respectively.

It is interesting that no clear relationship between land clearing in the sub-catchment and the O/E score could be obtained in this study. Townsend et al. (2004) showed that when land clearing in the catchment was over a threshold value of 50%, that there was a relationship between the amount of clearing and the O/E score of the Darwin-Daly genus level model.

Many other studies have demonstrated a reduction of macroinvertebrate diversity in relation to clearing in the catchment, however, these studies measured diversity in other ways than with the AUSRIVAS O/E score (e.g. Couceiro et al. 2007, Liess & Ohe 2005, Benstaed et al. 2003, Hogg & Norris 1991).

5.3.2 Response variable as percentage of horticulture and vegetation in sub-catchment

The genus O/E variable is the better variable in explaining the responsive variable percentage of vegetation and horticulture in sub-catchment, with about ninety nine percent strength of evidence. The fit of this linear model is with a correlation coefficient of 0.24 just outside the quarter goodness of fit, which is considered acceptable; statistically the fit is supported. It also indicates that most of the variation is not explained by land-use.

As discussed above, the O/E scores of the genus and family level model are correlated, with the correlation being stronger the larger the range of O/E scores covered. The genus O/E scores displayed a better statistical fit when explaining horticultural and vegetation land use in the sub-catchment, whereas land clearing was not well explained by either of the two AUSRIVAS models, possible because the relationship not linear (e.g. Townsend et al,. 2004).

When the two models were related to the land use in the catchment, the genus O/E was explainable through horticulture and vegetation, whereas the family O/E did not reveal a clear relationship to any of the land use types.

Overall, the genus level model allowed for some explanation of land use, whereas the family model was of limited use in this aspect.

Marshall et al (2006) found that in dryland streams in Southern Queensland species level presence / absence data - such as used in AUSRIVAS - were matched by the respective dataset merged on genus level with a 96 % Bray-Curtis similarity. The same dataset merged on the family level displayed a considerably lower Bray-Curtis similarity of 84 %. This confirms the expected higher sensitivity at a higher taxonomic resolution.

Marchant et al. (2006) also concluded that for low diversity, harsh dryland rivers in Queensland the family level analysis was superior from a cost/benefit analysis than respective data of lower taxonomic resolution.

Bearing in mind that Marchant et al (1995) found that small spatial scale studies, such as reaches, require lower taxonomic resolution to discriminate patterns, the genus level model is considered superior to the family level model for such small scale studies.

The genus level AUSRIVAS model is hence the more sensitive of the two models and is recommended, if resources allow. However, there was good correlation between the two AUSRIVAS models, especially when O/E scores are low due to significant site impairment. The family level model is therefore also acceptable for stream health assessment.

6. REFERENCES

Arscott DB, Jackson JK, Kratzer EB (2006): Role of rarity and taxonomic resolution in a regional and spatial analysis of stream macroinvertebrates. Journal of the North American Benthological Society, 25(4): 977-997.

Barlow A & Lamche G (2005): Development of an AUSRIVAS model for assessment of stream health in the Darwin-Daly region of the Northern Territory using family level data. Report to the National River Health Assessment Technical and Professional Advisory Committee; Canberra, 11p.

Benstead JP, Douglas MM, Pringle CM (2003): Relationships of stream invertebrate communities to deforestation in Eastern Madagascar. *Ecological Applications* 13(5): 1473-1490.

Breen P, Walsh C, Nichols S, Norris R, Metzeling L, Gooderham J (1999): Urban AUSRIVAS: An evaluation of the use of AUSRIVAS models for urban stream assessment. CRC for Freshwater Ecology Report No 5, LWRRDC Occasional Paper 12/99. ISSN 1320-0992; ISBN 0 642 26766 9.

Bureau of Rural Sciences (2002): Land Use Mapping at Catchment Scale: Principles, Procedures and Definitions. Bureau of Rural Sciences, Canberra.

Burnham KP & Anderson DR (2002): Model Selection and Multi-Model Inference: a Practical Information-Theoretic Approach. Springer-Verlag: New York.

Burnham KP & Anderson DR (2001): Kullback-Leibler information as a basis for strong inference in ecological studies. *Wildlife Research* 28, 111-119.

Burnham KP & Anderson DR (1998): Model Selection and Inference: a Practical Information-Theoretic Approach. Springer-Verlag: New York.

Couceiro SRM, Hamada N, Luz SLB, Forsberg BR, Pimentel TP (2007): Deforestation and sewage effects on aquatic macroinvertebrates in urban streans in Manaus, amazonas, Brazil. *Hydrobiologia* 575: 271-284.

Davies PE (2000): Development of a national river bioassessment system (AUSRIVAS) in Australia. In: Wright JF, Sutcliffe DW & Furse MT (eds): Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques. *Freshwater Biology,* pp 113-124.

Davies, PE (1994). National River Processes and Management Program. Monitoring River Health Initiative. River Bioassessment Manual, Version 1.0. Department of Environment, Sport and Territories, Land and Water Resources Research and Development Corporation, Commonwealth Environment Protection Agency.

Dostine P. (2004): A predictive model for assessment of stream health in the Darwin –Daly region of the Northern Territory using mostly genus level data. Report 42/2004D, Department of Infrastructure, Planning & Environment, Natural Resource Management, 27p.

Dostine P (2002): Australia-wide Assessment of River Health – Northern Territory Program – Final Report. Report 07/2002, Department of Infrastructure, Planning & Environment, Natural Resources Division, 92p.

Fukuda Y & Townsend S (2006): Dry Season Water Quality resource Condition Targets for the Rivers and Streams of the Darwin-Litchfield-Bynoe Region. Report 15/2006D. Department of Natural Resources, Emvironment and the Arts, Darwin NT.

Hogg ID & Norris RH (1991): Effects of Runoff from Land Clearing and Urban Development on the Distribution and Abundance of Macroinvertebrates on Pool Areas of a River. *Australian Journal of Marine and Freshwater Research* 42, 507-518.

Lamche, G (2007): The Darwin-Daly regional AUSRIVAS models – Northern Territory. User Manual. Report 06/2007D. Aquatic Health Unit, Environment Protection Agency Program, Department of Natural Resources, Environment and the Arts, Darwin. ISBN 1-920772-44-8.

Marchant R, Barmuta LA, Chessman BC (1995): Influence of Sample Quantification and Taxonomic Resolution on the Ordination of Macroinvertebrate Communities from Running Waters in Victoria, Australia. *Marine and Freshwater Research* 46: 501-506.

Marshall JC, Steward AL, Harch BD (2006): Taxonomic resolution and quantification of freshwater macroinvertebrate samples from an Australian dryland river: the benefits and costs of using species abundance data. *Hydrobiologia* 572: 171-194.

Nichols SJ, Robinson WA, Norris RH (2006): Sample variability influences on the precision of predictive bioassessment. *Hydrobiologia* 572: 215-233.

Simpson JC & Norris RH (2000): Biological assessment of river quality: development of AUSRIVAS models and outputs. In: Wright JF, Sutcliffe DW & Furse MT (eds): Assessing the Biological Quality of Fresh Waters: RIVPACS and Other Techniques. *Freshwater Biology*, pp 125-142.

Sloane PI & Norris RH (2003): Relationship of AUSRIVAS-based macroinvertebrate predictive model outputs to a metal pollution gradient. *Journal North American Benthological Society* 22(3): 457-471.

Townsend, S.A., Dostine, P.L., Dixon, I., Karfs, R. and Douglas, M.M. (2004). The response of stream macroinvertebrate communities to catchment clearing and riparian condition in the Darwin region (tropical Australia). In 'Proceedings of the 4th Australian Stream Management Conference'. (Ed. Rutherford, ID, Wiszniewski, I, Askey-Doran, M., and Glazik, R Department of Primary Industries, Water and Environment.); pp 618-622.

Waite IR, Helihy AT, Larsen DP, Urquhart NS, Klemm DJ (2004): The effects of macroinvertebrate taxonomic resolution in large landscape bioassessments: an example from the Mid-Atlantic Highlands U.S.A. *Freshwater Biology* 49 (4): 474-489.

Wright JF (1995): Development and use of a system for predictingthe macroinvertebrate fauna in flowing waters. *Australian Journal of Ecology* 20: 181-197.