A review of load tools available for calculating pollutant exports to the Great Barrier Reef lagoon: A case study of varying catchment areas

Lewis, S.E.^{1,2}, Z.T. Bainbridge^{1,2} and J.E. Brodie^{1,2}

¹ Australian Centre for Tropical Freshwater Research, James Cook University, Townsville 4811 Australia. ² Marine and Tropical Sciences Research Facility Email: Stephen.Lewis@jcu.edu.au

Keywords: Load tools, Load methods, GBR catchment loads, target setting

EXTENDED ABSTRACT

Several load methods and software programs have been developed to calculate sediment, nutrient and other pollutant exports from waterways of the Great Barrier Reef (GBR) catchments. These different methods can produce large discrepancies in the calculation of catchment loads. Such discrepancies reduce the confidence of these methods for application within the Reef Water Quality Protection Plan process, such as the setting of end-of-river load targets as well as the comparison to modelled outputs, namely the SedNet and ANNEX models (e.g. Brodie et al. 2003). We present a case study based on intensive monitoring data collected over a range of spatial (paddock, a Dry Tropics sub-catchment and a large Dry Tropics end-of-river catchment) and temporal (hourly - yearly) scales. We simulate changes in load calculations of total suspended solids based on the selected removal of monitoring data from intervals throughout the flow hydrograph. We attempt to quantify errors in load calculations based on the different load tools, as well as investigate optimal load methods and ideal sampling frequencies over these different catchment areas. Three software programs were used to calculate loads including WO Loads Tool, Brolga and GUMLEAF. We found that all three software programs provided suitable methods to calculate loads for the catchments of the GBR where continuous flow and concentration data were available. The linear interpolation (and associated inter sample mean) methods were optimal in the Brolga and Loads Tool program while the flow regime estimators provided ideal load estimates in the GUMLEAF program. Our findings suggest that six samples evenly-spaced over the flow hydrograph (e.g. 2-3 samples on rise, 1 on peak and 2-3 on the falling limb) will provide reliable load estimations (within 10% of our best estimate) at the paddock scale provided that the optimal methods are used. We recommend that at least daily (but up to 4-5 samples per day in

catchments with very high TSS concentrations on the rising limb such as the Bowen sub-catchment) sampling is suitable for load estimations at the sub-catchment scale. One sample collected every two days is an adequate sampling frequency for load calculations of larger catchments of the GBR such as the Burdekin River. We note that researchers need to account for the uncertainty in all load estimates before the significance of longterm trends can be analysed.

1. INTRODUCTION

Knowledge of the loads of materials transported through waterways is critical in water quality studies to identify pollutants of greatest concern, to quantify changes in water quality due to incatchment actions, to set water quality targets and to assess the validity of predictive models. Loads are calculated by using the continuous flow volume of the waterway (commonly measured from hourly to daily) in combination with the concentration of a particular material to calculate the total mass exported through the sampled point Several equations have been of the stream. formulated to calculate loads which have varying degrees of complexity, ranging from the relatively simple flow by concentration technique to more complex ratio equations (Letcher et al. 1999). Previously, most calculations of loads were complex, tedious and time-consuming. Recently, a range of software programs have been developed (WQ Loads Tool, Brolga and GUMLEAF) to calculate loads more easily. These programs are now utilised by many scientists and natural resource managers. While these programs have greatly increased productivity, users still must select the optimal method to calculate loads for their particular dataset and understand the limitations and errors inherent in the load estimations produced.

Selection of a load method is strongly dependent on the concentration/stream flow data available, the hydrological characteristics of the waterway and the desired accuracy required (Letcher et al. 1999, Fox et al. 2005). The majority of load equations have been developed for catchments in the Northern Hemisphere and many are not applicable for Australian streams. Research efforts on Australian streams have been concentrated in the temperate environments in south-eastern Queensland, New South Wales and Victoria (Letcher et al. 1999, Fox et al. 2005). The optimal method(s) to calculate loads for the catchments of the Great Barrier Reef (GBR) in the Wet and Dry Tropics of Queensland has not been thoroughly investigated. Moreover, the load methods (and software programs) need to be assessed over different catchment areas and over different sampling regimes to understand the optimal sampling intervals and the best available method that will provide the desired accuracy and precision. We employ four datasets generated from three GBR catchments, with areas ranging from 2 Ha to 130,000 km², to investigate the optimal load method for different sampling scenarios and over different temporal scales (two days to one year). We use total suspended solids

(TSS) to represent loads for the particulate materials transported through the waterways.

2. BACKGROUND

The knowledge of catchment loads along E Australia has never been as important (or received such attention) since the introduction of the Reef Water Quality Protection Plan for the GBR in 2003 (Anon, 2003). Previously sediment and nutrient loads were calculated to assess the progradation of the inner shelf of the GBR (Belperio, 1983) and to construct nutrient budgets for the GBR lagoon (Furnas et al. 1995). Recently, the major motive for load studies has been to estimate material loading to the GBR and several models have been developed to estimate catchment exports (e.g. Brodie et al. 2003). A plethora of water quality data have also been collected in the GBR catchments (e.g. Mitchell et al. 2006, Bainbridge et al. 2006) and so catchment loads can be calculated and compared to the model estimates (e.g. Mitchell et al. 2006; Sherman et al. 2007). This comparison provides valuable data to assist with the further development and improvement of the models so that simulations can be made with increased confidence and the influence of changes in catchment condition and land use on material loads can be evaluated.

There are several different equations within three general techniques to calculate loads including mean-based estimators (e.g. flow by concentration, linear interpolation), regression estimators (e.g. rating curves) and ratio estimators (e.g. Beale's Ratio, Kendall's Ratio) (Letcher et al. 1999). These equations produce varying degrees of bias and precision and there is no single optimal technique for load calculations; the most reliable technique depends on the objectives of the study (e.g. accuracy and precision required), and the flow and water quality data (e.g. sampling interval, sampling focus) available. Most studies suggest that a flow stratified sampling approach that biases towards high flow events should be adopted for Australian streams (Letcher et al. 1999, Fox et al. 2005). In particular, the linear interpolation method appears to be the most suitable for Australian streams provided that continuous flow data are available and that a flow stratified sampling approach is conducted (Letcher et al. 1999, Fox et al. 2005). Some studies have postulated that the collection of five to eight samples across the flow hydrograph will provide reliable load estimates (NR&W, 2006); however, the sampling frequency required at different catchment scales and different hydrological regimes is currently unresolved.

Table 1. The benefits and limitations of the three software programs availabl	e to calculate loads.
---	-----------------------

Software program	Benefits	Limitations
GUMLEAF v0.1alpha - Australian Centre for Environmetrics (Tan et al. 2005)	 Provides several (22) methods to calculate loads. Provides an estimation of load uncertainty. Easy to use and provides rapid load estimations. 	 Not useful for smaller catchment sizes where sub-daily sampling is required. Current version only calculates loads using daily intervals. Cannot view flow hydrographs of flow/concentration data. Data stored in Microsoft Excel files and not in a specific database.
Brolga v2.11 - Queensland Department of Natural Resources and Water (NRW, 2007)	 Loads can be calculated at any time interval up to yearly. Program suitable for use at different catchment sizes. Flow and concentration data are stored into Microsoft Access database. Outputs graphs of flow and concentration data can be modified and used in reports. Program also shows in graph form the assumptions made by the three load methods used. 	 A little more time-consuming to enter data to calculate loads. Only three load methods to choose from. Does not provide an estimation of load uncertainty.
Loads Tool v1.0.1- NAP product (Marsh et al. 2006)	 Provides nine methods to calculate loads. Simple to use by importing flow and concentration data into the program from csv. files. Provides rapid load estimations. Program designed for both long-term and short-term (event) load calculations at different catchment sizes. Provides an EMC for the load period. 	 Graphs produced in the program can not be used in reports. Data stored in Microsoft Excel files and not in a specific database. Does not provide an estimation of load uncertainty.

Three software programs have recently been developed to calculate catchment loads including Generator for Uncertainty Measures and Load Estimates using Alternative Formulae (GUMLEAF version 0.1 alpha: Tan et al. 2005), Brolga (version 2.11: Natural Resources and Water [NR&W], 2007) and Loads Tool (version 1.0.1: Marsh et al. 2006). The programs are all simple to use and provide rapid load calculations, although they also have some minor limitations (Table 1). In this study, we apply these programs to produce load estimations for three different sized catchments of the GBR. We investigate the optimal load method at each catchment scale and also examine the minimum sampling efforts required to produce reliable load estimations.

3. METHODS

3.1. Paddock-scale samples

An excellent flow and water quality dataset has been collected from a banana farm situated within the Wet Tropics near the township of Tully (Fig. 1A). The watershed area is approximately 2 Ha and the data have been reported by Faithful et al. (2007). Continuous flow data were recorded by a gauge installed by NR&W while water samples were collected using an ISCO autosampler. The autosampler was set to collect samples at 10 min intervals for a duration of 2 hours and then at 30 min intervals for the following 3 hours. Samples were then collected at hourly intervals. This sampling design targets the initial first flush and high flow intervals where the highest concentrations and

fluxes of materials are transported. The data used in this study was collected during a runoff event between the 11th-13th of January 2006. TSS analysis was conducted at the Australian Centre for Tropical Freshwater Research (Faithful et al. 2007). Five scenarios were devised to simulate different sampling intervals and the data were imported into the software programs to calculate loads using all available formulae. The five scenarios for the paddock-scale catchment included: 1. samples collected at the autosampler setting; 2. samples collected at hourly intervals; 3. samples collected at two-hourly intervals; 4. sampling missed on rising stage flow; and 5. six samples collected over the flow hydrograph.

3.2. Sub-catchment scale samples

The sub-catchment scale dataset was taken from the Bowen River at the Myuna Station containing a catchment area of $7,200 \text{ km}^2$. This subcatchment is situated within the Burdekin River catchment in the Dry Tropics. A flow gauging station (no. 120205a: Bowen River @ Myuna) on the site is operated by NR&W which collected hourly data (m³/second) while an ISCO autosampler was used to collect water samples. The autosampler collected samples on a stagebased procedure (with a bias towards the rising limb) where height, time and the amount of bottles remaining in the carousel are taken into consideration so that samples were taken over the entire flow hydrograph (R. Keen pers comm. TSS samples were analysed by the 2007). Queensland Health Scientific Services, Brisbane (Bainbridge et al. 2006). Data collected during a flow event which occurred from the 27th-30th January 2006 were used for load analysis. The first flush samples from the Bowen River contained extremely high TSS concentrations (peak of 14,000 mg/L) before falling below 2,000 mg/L within twelve hours of flow (Fig. 1B). The high TSS concentrations during the first flush are a typical feature of GBR catchments (Bainbridge et al. 2006), although the Bowen River catchment contains considerably higher TSS concentrations on the rising limb. The five scenarios devised for this sub-catchment included: 1. all samples collected by autosampler; 2. samples collected at five hourly intervals; 3. samples collected at daily intervals; 4. sampling missed on rising stage flow; and 5. six samples collected over the flow hydrograph.

3.3. End-of-catchment scale samples

The Burdekin River is one of the largest of the GBR catchments (130,000 km²) and is situated within the Dry Tropics Region. Mean annual discharge for the Burdekin River is 8.43 million ML, although river flow is highly variable (Bainbridge et al. 2006). An extensive end of river monitoring program was conducted by the Australian Institute of Marine Science (AIMS) from 1987 to 2000 (Mitchell et al. 2006). Samples were collected at intervals of up to twice-daily (morning and afternoon) during large flow events and TSS samples were analysed at the AIMS laboratory (Mitchell et al. 2006). Flow data is recorded by the NR&W gauge (no. 120006b: Burdekin River @ Clare) located approximately 20 km upstream from the sampling site (note that there are no major confluences between these points). Data collected in the 1996/97 (Fig. 1C) and 1999/00 (Fig. 1D) water years (1st October to 30th September) were used to calculate annual loads. While both water years produced above average annual discharge (8.66 and 13.32 million ML, respectively) and contained several flow events, these water years produced two distinct event flow hydrographs; the 1996/97 water year is characterised by four large discharge pulses while the 1999/00 water year has only one large discharge pulse (Figs. 1C and 1D). The five scenarios devised for this catchment for both water years included: 1. samples collected as sampled by AIMS; 2. samples collected at daily intervals; 3. samples collected every two days; 4. sampling missed on rising stage flow; and 5. limited samples (eleven samples 1996/97 and six samples 1999/00) spaced over the flow hydrograph.



Figure 1. Flow hydrographs and TSS concentrations (mg/L) of the paddock-scale (A), sub-catchment scale (B) and end-of-catchment scale (C: 1996/97 water year; D: 1999/00 water year) of the data used for load calculations.

3.4. Load analysis

As it is not possible to know the 'true load' for any of these datasets, we have used the linear interpolation equation from the Brolga program to compare the variability of the other load methods as a 'best estimate'. The linear interpolation method has been used as the 'true load' in previous studies to assess load estimates (Letcher et al., 1999) and has been postulated as the best technique to calculate loads where continuous data are available (Fox et al. 2005). The three software programs combined provide up to 34 load calculations and also allow for comparisons to be made with similar methods used in the programs (e.g. Beale Ratio, linear interpolation). We note that the GUMLEAF program in its current form (version 0.1 alpha) does not support event flow estimates at the sub-daily (to hourly) sampling intervals and so was not used in the paddock scale data. We also note that there may be some variability in the load estimates in the reduced sampling strategies based on the selection of the first sample, although we believe that this exercise still highlights the variability encountered in load estimates based on different sampling strategies. Throughout the remainder of the document the five scenarios will be abbreviated as S1 to S5.

4. RESULTS AND DISCUSSION

4.1. Paddock-scale sampling

Three methods provided consistent TSS load estimates (within 10%) over the five scenarios at the paddock-scale catchment including linear interpolation, inter sample mean concentration and inter sample mean concentration using mean flow (all three methods from the WQ Loads Tool). The Brolga linear interpolation method produced load estimates within 15% across the five scenarios. In general, both the two hourly sampling (S3) and samples not taken on rising stage flow (S4) scenarios resulted in lower load estimates while several methods provided accurate load estimates where sampling was spaced evenly over the flow hydrograph (S5). However, there was high variability between methods within scenarios (in S1 the methods ranged from 6.8 kg - concentration power curve fit method (WO Loads Tool) to 1000 kg - average load: WQ Loads Tool) and most individual methods were also highly variable across scenarios (e.g. the average load method ranged from 350 in S3 to 1,720 kg in S5). Therefore most methods were considered unsuitable for load estimates at the paddock-scale with the exception of the linear interpolation and inter sample mean flow methods.

Our analysis suggest that the best methods available to calculate loads at the paddock scale were linear interpolation (both Brolga and Loads Tool), inter sample mean concentration and inter sample mean concentration using mean flow (Loads Tool). Interestingly, there was a 20% difference in the loads estimated by the two linear interpolation methods in the Brolga and Loads Tool programs. This difference suggests that the lowest error possible for 'best estimates' at the paddock scale is approximately 20% across the software programs at this sampling strategy; this discrepancy needs further investigation. The data show that collecting six samples spaced evenly over the hydrograph (includes two-three samples on the rise, one on the peak and two-three samples on the fall) is suitable for reliable (within

10-15% of best estimate) load estimates at the paddock scale provided that the best methods are utilised.

4.2. Sub-catchment scale sampling

Most methods in S1 for the Bowen River subcatchment provided TSS load estimates within 10% of the linear interpolation estimates (Brolga and Loads Tool). This finding suggests that the autosampler setting at this site was adequate for reliable load calculations and that a range of methods will produce reliable (within 10% of best estimate) results at this sampling frequency. However, where the sampling frequency was reduced, all methods produced considerably lower load estimates of around 20% (S2: 5 hourly sampling), 60% (S3: daily sampling) and 80% (S4: rising stage samples missed). This result highlights that careful sampling of this catchment is required to produce accurate load estimates. The two linear interpolation methods and the inter sample mean concentration methods provided accurate loads (within 10%) where six samples were taken evenly over the hydrograph (S5) which incorporated the very high TSS concentrations on the rising stage of the hydrograph. The GUMLEAF program was used for the daily sampling scenario (S3) and the three flow regime stratified methods (flow weighted mean concentration, simple ratio estimator and Beale's Ratio estimator) provided loads within 10% of our 'best estimated load'.

Similarly to the paddock scale estimates, we recommend the use of the linear interpolation, inter sample mean concentration and inter sample mean concentration using mean flow methods. In this larger catchment, the direct comparison between the two linear interpolation methods was within 5% of each other across the five scenarios. We suggest that high frequency sampling (similar to the autosampler setting) is required to produce reliable load estimates (within 10-15% of the best estimate) for the Bowen sub-catchment, which is peak characterised by very high TSS concentrations on the rising limb. Where peak TSS concentrations are not as pronounced, daily sampling would probably be sufficient for reasonable load estimates. The collection of six samples spaced over the hydrograph provided reasonable load estimates (within 10%) in the Bowen sub-catchment for the better methods but this sampling design should be concentrated particularly on the rise and peak of the flow hydrograph. It is emphasised, however, that the load methods for this scenario (S5) produced estimates which ranged from 27,000 (flow weighted concentration) to 270,000 tonnes (average load), an order of magnitude difference. Caution is required when loads are calculated using a limited number of samples.

4.3. End-of-catchment sampling

Most of the load estimations for S1 for both the 1996/97 and 1999/00 water years produced consistent (within 10%) results, suggesting a reliable sampling strategy was employed. The GUMLEAF program was used to calculate annual TSS loads for scenarios 2 to 5. The 'better load methods' (Brolga: linear interpolation; WQ Loads Tool: linear interpolation, inter sample mean concentration: GUMLEAF: flow regime stratified estimators) all produced consistent results (within 10%) over scenarios 1, 2 and 3 in both water years. This finding indicates that sampling the Burdekin River every second day during flow events should provide reliable load estimates. The comparison of S4 to the 'best estimated load' for both water years provided some variable results. The better methods typically underestimated loads by ~20% in the 1996/97 water year which was characterised by droughtbreaking floods (i.e. higher TSS concentrations in the 'first flush' which was not sampled in this scenario) while these same methods provided loads closer (within 10%) to our best estimate in the 1999/00 water year where the rising stage samples were missed (S4). The better methods in the S5 simulation (eleven samples over hydrograph) for the 1996/97 water vear consistently overestimated loads by ~30%. This overestimation was probably due to the preferential selection of samples on the rising limb of the hydrograph coinciding with highest TSS concentrations, which thus biased the overall load estimate. In contrast, the better methods in the Brolga and Loads Tool programs underestimated loads by ~30% in S5 for the 1999/00 water year, while the recommended methods in GUMLEAF overestimated loads by ~50%.

There was excellent agreement between the linear interpolation methods in Brolga and the WQ Loads Tool in all Burdekin River scenarios and reasonable comparisons between the Beale Ratio (Loads Tool) and the annual Beale's Ratio estimator (GUMLEAF) with the exception of S4 (~30%) in the 1996/97 water year and in S5 $(\sim 40\%)$ in the 1999/00 water year. We recommend both linear interpolation methods in the Brolga and Loads Tool programs, inter sample concentration, inter sample mean mean concentration using mean flow (WQ Loads Tool), and the flow regime stratified estimators (GUMLEAF) as the optimal techniques for

reliable load estimations. We suggest, for large Dry Tropics catchments such as the Burdekin River, that sampling during event flows should be conducted on a minimum of every second day. The data from this study show that the sampling strategy of six-eight samples spaced over the hydrograph is not sufficient for a catchment of this size.

4.4. Uncertainty analysis

The only program currently available to quantify the uncertainty of load estimates is GUMLEAF program which uses a Monte-Carlo approach (Etchells et al. 2005). An analysis of this uncertainty, however, is beyond the space limits of this paper.

5. CONCLUSIONS

We used up to 34 different methods in three software programs to calculate loads over three different catchment areas of the GBR to examine the optimal load method and the suitable sampling frequencies over these catchment scales. We found that all software programs provided suitable load methods for the catchments of the GBR (where continuous data were available). We note that in its current form, the GUMLEAF program can not be used at smaller catchment scales. The optimal methods were:

- Brolga: linear interpolation;
- WQ Loads Tool: linear interpolation, inter sample mean concentration, inter sample mean concentration using mean flow;
- GUMLEAF: flow regime stratified flow weighted mean concentration estimator (method # 19), flow regime stratified simple ratio estimator (method # 20), flow regime stratified Kendall's Ratio estimator (method # 21) and flow regime stratified Beale's Ratio estimator (method # 22).

The minimum sampling frequencies recommended for the different catchment areas were:

- Paddock scale: six samples evenly spaced over the hydrograph (at least two on rising limb);
- Sub-catchment scale: daily sampling (although for catchments with very high material concentrations on the rising limb such as the Bowen River, 4-5 samples per day may be required);
- Large end-of-catchment scale (e.g. Burdekin River): one sample collected every two days.

We believe that this exercise will help in the selection of a suitable load method for specific catchments, and encourage similar studies to build on these findings. Researchers also need to assess the potential errors and uncertainty inherent in load estimates. Only once this assessment is understood (or even quantified) can the significance of long-term trends in water quality be confidently deduced through the use of catchment loads.

6. ACKNOWLEDGEMENTS

We thank the CSIRO Water for a Healthy Country GBR collaborative program with JCU for supporting this research. We thank John Faithful, Miles Furnas and Alan Mitchell for use of concentration data in this study. We acknowledge Burdekin Dry Tropics NRM for providing funding for the Bowen River results and thank CSIRO Land and Water for maintaining the autosampler at the Myuna site. The NRM hydrographers (Ayr) provided the relevant flow data. Many thanks to the managers of the GUMLEAF, Brolga and WQ Loads Tool for supplying the software programs. We thank MTSRF for support in this project.

7. REFERENCES

- Anon. (2003), Reef Water Quality Protection Plan; for catchments adjacent to the Great Barrier Reef World Heritage Area, Queensland Department of Premier and Cabinet, Brisbane.
- Bainbridge, Z., Lewis, S., Brodie, J., et al. (2006), Monitoring of sediments and nutrients in the Burdekin Dry Tropics Region: 2005/06 wet season, ACTFR Report 06/13, JCU, Townsville.
- Belperio, A.P. (1983), Late Quaternary terrigenous sedimentation in the Great Barrier Reef, in: Baker, J.T. Carter, R.M. Sammarco, P.W. Stark, K.P. (eds) *Inaugural Great Barrier Reef Conference*, James Cook University Press, Townsville, p 71-77.
- Brodie, J. McKergow, L.A. Prosser, I.P. Furnas, M. Hughes, A.O. Hunter, H. (2003), Sources of sediment and nutrient exports to the Great Barrier Reef World Heritage Area. ACTFR Report 03/11, JCU, Townsville.
- Faithful, J.W. Brodie, J. Hooper, A. Leahy, P. Henry, G. Finlayson, W. Green, D. (2007),

Plot-scale runoff of nutrients and sediment under varying management regimes of a banana and cane farm in the Wet Tropics, Queensland. ACTFR Report 05/03, JCU, Townsville.

- Fox, D.R. Etchells, T. Tan, K.S. (2005), *Protocols* for the optimal measurement of nutrient loads, A report to West Gippsland Catchment Management Authority, Australian Centre for Environmetrics, July 2005.
- Furnas, M. Mitchell, A.W. Skuza, M. (1995), Nitrogen and phosphorus budgets for the central Great Barrier Reef shelf, Great Barrier Reef Marine Park Authority Research Publication no. 36, Great Barrier Reef Marine Park Authority, Townsville.
- Letcher, R.A. Jakeman, A.J. Merritt, W.S. McKee, L.J. Eyre, B.D. Baginska, B. (1999), *Review of techniques to estimate catchment exports*, Environmental Protection Authority Technical Report, EPA, Sydney. ISBN: 0731327098.
- Marsh, N. Steven, A. Tennakoon, S. Arene, S. Farthing, B. Fox, D. (2006), Loads Tool v1.0.1, QNRM06085.
- Mitchell, A. Furnas, M. De'ath, G. Brodie, J. Lewis, S. (2006), A report into the water quality condition of the Burdekin River and surrounds based on the AIMS end-ofcatchment sampling program, ACTFR Report 06/06, JCU, Townsville.
- NR&W, (2006), Manual sampling guidelines, Department of Natural Resources and Water, The State of Queensland.
- NR&W. (2007), Brolga version 2.11, DNRW.
- Sherman, B., Brodie, J., Cogle, L., Carroll, C. (2007), *The appropriate use of catchment models for water quality target setting and land use management*, in: IAHS Symposium on Water Quality and Sediment Behaviour of the Future: Predictions for the 21st Century, 2007.
- Tan, K.S. Etchell, T. Fox, D.R. (2005), GUMLEAF v0.1alpha: Generator for Uncertainty Measures and Load Estimates using Alternative Formulae, Australian Centre for Environmetrics, University of Melbourne.