Great Barrier Reef Source modelling: Assessing hillslope erosion modelling performance at paddock scale experimental sites

C. Dougall * and G. McCloskey *

* Department of Natural Resources and Mines, Queensland

Email: Cameron.Dougall@dnrm.qld.gov.au

Abstract: Catchment modelling is used as one of multiple lines of evidence to report on progress towards Great Barrier Reef (GBR) Water Quality Protection Plan (Reef Plan) water quality targets. Progress is monitored via the Paddock to Reef (P2R) Integrated Monitoring, Modelling and Reporting Program. Within the P2R program, the Source modelling framework is used to construct GBR specific models (McCloskey 2017) and the overarching purpose of the modelling is to assess the impact of land management changes across the GBR against Reef Plan targets.

To optimise predictive capability and thereby help facilitate better targeting, the modeller is faced with two common problems. Firstly, assessing the quality of model inputs and outputs and secondly, where possible, improving modelling performance. While observation datasets such as flow are often adequate, water quality data are scarcer, sporadically sampled, and frequently well downstream of their source. To further improve the water quality model validation data pool, the modelling team have attempted to tabulate all relevant datasets that can be used in model evaluation. However, utilising all available datasets to assess and improve performance can be an arduous task. As such, small catchment studies can initially be overlooked as a validation source in large “lumped” catchment scale modelling studies due to scale and the inherent spatial and temporal mismatch issues.

As a small illustrative case study, we firstly designed a simple methodology to disaggregate the lumped modelled hillslope loads for a grazing landuse in a subcatchment. We then discuss the benefit of paddock scale research sites for informing the GBR Source modelling.

In the GBR Source framework, spatially variable Revised Universal Soil Loss Equation (RUSLE) was used to generate the hillslope sediment loads for rangelands. Sediment loads are lumped and reported for a single landuse (typically1 km² – 30km²) for a given sub-catchment at a daily timestep. However the modelled daily loads are initially generated at a finer scale, typically, 30m by 30m - 900 m². As the individual grid calculations are not recorded at model runtime, they needed to be recorded prior to aggregation in the model. This enabled evaluations between GBR Source modelled loads and the much smaller catchment study data to be made at appropriate scales.

The disaggregation and assessment of modelled loads outlined in this paper facilitated small to large scale hillslope erosion comparisons. For the two case study sites, erosion and cover are shown to be well predicted at the average annual scale and areas for further improvement and investigation have been highlighted. We recommend further sites across the GBR be assessed.

Validating disaggregated modelled data against plot scale experimental data, provides an alternative data set for validation previously not utilised. The approach shows a lot of potential to improve confidence in modelled sediment generation predictions.

Keywords: GBR Source model, GBR water quality, fine sediment
Dougall and McCloskey, Great Barrier Reef Source Catchment’s modelling: Assessing hillslope erosion modelling performance at paddock scale experimental sites

1. INTRODUCTION

Catchment modelling is used as one of multiple lines of evidence to report on progress towards Great Barrier Reef (GBR) Water Quality Protection Plan (Reef Plan) water quality targets. Progress is monitored via the Paddock to Reef (P2R) Integrated Monitoring, Modelling and Reporting Program (Carroll et al. 2013). The Reef Plan water quality targets are set for the whole of the GBR, which encompasses six contributing NRM regions: Cape York, Wet Tropics, Burdekin, Mackay Whitsunday, Fitzroy and the Burnett Mary. Within the P2R program, the Source modelling framework is used to construct GBR specific models (McCloskey 2017) and the overarching purpose of the modeling is to assess the impact of land management changes across the GBR against Reef Plan targets.

The GBR Source models serve a number of purposes one of which is to simulate accurate average annual loads at large catchment and sub-catchments scales. The models also have to mimic temporal generation and delivery processes and simulate the relative impact of differences in land management practice. Reef Plan requires the modelling of many water quality constituents, including fine sediment (FS), nutrients and pesticides. Given these requirements, the GBR Source models have been configured to model 10 constituents, at a daily timestep over a 28 year time frame, with a large array of process representation (McCloskey 2017).

To optimise predictive capability and thereby help facilitate better targeting of on ground practices, the modeller is faced with two common problems. Firstly, assessing the quality of model inputs and outputs and then, secondly, where possible, improving modelling performance. While observation datasets such as flow are often adequate, water quality data are more scarce, sporadically sampled, regularly well downstream of their source (Dougall and Carroll 2013). To further improve the water quality model validation data pool, to the modelling team are tabulating all relevant datasets that can be used in model evaluation. To help simplify communication, prioritise effort and assess data uncertainty, we have used the classification system of Seibert and McDonnell (2002) where datasets are divided into “hard” and “soft” classes (Table 1).

<table>
<thead>
<tr>
<th>Data Type</th>
<th>Type</th>
<th>Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hydrology</td>
<td>Rainfall</td>
<td>Hard</td>
</tr>
<tr>
<td>Constituent</td>
<td>WQ sampling</td>
<td>Hard</td>
</tr>
<tr>
<td></td>
<td>Stream flow</td>
<td></td>
</tr>
<tr>
<td>Research Sites</td>
<td>All + Contextual</td>
<td>Both</td>
</tr>
<tr>
<td>Proxys</td>
<td>Retreat rates</td>
<td>Soft</td>
</tr>
<tr>
<td></td>
<td>Sediment loads</td>
<td>Soft</td>
</tr>
<tr>
<td></td>
<td>Deposition rates</td>
<td>Soft</td>
</tr>
<tr>
<td></td>
<td>Geochemical, etc</td>
<td>Soft</td>
</tr>
<tr>
<td>Modelling</td>
<td>Paddock</td>
<td>Soft</td>
</tr>
<tr>
<td></td>
<td>Catchment</td>
<td>Soft</td>
</tr>
</tbody>
</table>

Briefly, “hard” data is data that is more directly measured, has low uncertainty, and is characterised by little data transformation or extrapolation (i.e. water quality samples, and well sampled loads). “Soft” data has much larger uncertainty bounds and is characterised by large amounts of data transformation or extrapolation (i.e. sediment fingerprinting (Koiter et al. 2013)), and can include expert opinion or local knowledge. Although inherently subjective, the definition can be conceptualised as lying on a numerical continuum, with zero being completely false and one being completely true. Importantly both dataset types contain information on the behaviour of the system and should be included in the modelling process if at all possible (Yen et al. 2016).

Nonetheless, utilising all available datasets to assess and improve performance can be an arduous task. As such small catchment studies can initially be overlooked as a validation data sources in large “lumped” catchment scale modelling studies. This is due to the inherent spatial and temporal mismatch in scale (Baffaut et al. 2015). As such, large scale catchment modelling projects traditionally focus on model evaluation at well studied downstream water quality sites, where reliable loads can be constructed. Yet small catchment scale studies in the GBR can provide a wealth of information about processes, particularly hydrology, the erosion of fine sediment (FS), and the impact of management change on erosion rates. Allocating scarce modelling resources to the investigation of FS has a particular advantage, in that often it is the best sampled of the constituents and correlates well with particulate phosphorus and to a lesser extent particulate nitrogen. Eroded sediment in GBR catchments is often classified into hillslope, gully, and streambank sources. Although hillslope sediment is reported as being a low contributor to GBR loads (Wilkinson et al. 2013), they are often enriched with particulate nutrients compared to other sources. To date there have been few analyses on the performance of catchment scale hillslope erosion modelling in the GBR (Silburn 2011). Here as a small illustrative case study, we firstly designed a methodology to disaggregate the lumped modelled hillslope loads and secondly report on the benefit of paddock scale research sites for informing the GBR Source modelling.

This paper demonstrates the benefit of utilising small catchment scale studies as a further data source in the challenging task of validating and calibrating GBR Source models. Encouragingly, reasonable hillslope erosion model performance is evident for the small case study catchments. Importantly, ground cover is also being accurately populated at small scales, suggesting good model representation of a key management practice change element.
2. METHOD

2.1. Study Area and Site Selection

The GBR catchment area is 423,000 km², with a north to south distance of approximately 1,800 km. The catchment experiences a typical sub-tropical climate with humid, wet summers and mild, dry winters. Average yearly rainfall in the catchment ranges from over 3,000 mm in north-eastern parts to less than 500 mm in south-western areas; however totals can be highly variable due to climatic drivers such as the El Niño Southern Oscillation (ENSO). Major landuses are grazing (~75% of the catchment), conservation (~5%), dryland cropping (~2%), and sugarcane (~1%).

Within the GBR catchments, there is a relatively short history of paddock scale constituent loss studies. A comprehensive list for GBR catchments was tabulated via interrogation of a database of catchment studies (http://howleaky.net/index.php/library), combined with local knowledge of missing sites. Some notable small grazing studies include the Brigalow catchment study (Cowie, Thornton, and Radford 2007), Virginia Park (Bartley et al. 2014), and Wambiana (O'Reagain et al. 2005). Cropping studies include sugarcane monitoring sites in the Mackay region, banana trials at South Johnstone, and grazings monitoring in the Fitzroy.

Of the initial 20 sites considered, nine sites were selected for assessing hillslope erosion model performance as they met the criteria of being: well documented (including soil erosion and hydrology load estimates), and datasets greater than two years in duration. To date we have completed analysis on four main sites. However, for brevity the Virginia Park and Brigalow catchment studies are only reported in this paper.

The Virginia Park study (Bartley et al. 2014) is located in the Burdekin basin, and is instrumented with three hillslope flumes (0.002-0.01 km²) and gauged at the larger nested catchment scale (14 km²). The record is much shorter than the Brigalow study, with semi-continuous data available on erosion rates and ground cover for the period 2001 to 2011.

The Brigalow Catchment Study (Cowie, Thornton, and Radford 2007) is a paired, calibrated catchment study in the Fitzroy basin. It is instrumented on identical soil types with three small catchment flumes (0.12-0.17 km²) monitored since 1965. The sites represent the three major regional landuses: native forest, cropping and grazing. Continuous data is available on erosion rates for the period 1987-2010 (Elledge and Thornton 2017).

2.2. Model Description and Computation

Globally, surface erosion (rill and interrill) is commonly calculated using the Universal Soil Loss Equation (USLE) and its derivative the Revised Universal Soil Loss Equation (RUSLE) (Renard et al. 1997). Observations of soil erosion on research plots (defined as 22.1m long, by 4m wide) were used to derive the model relationships. The USLE predicts average annual erosion at the plot scale using six factors multiplied together in the following linear form.

\[ A = R \times K \times L \times S \times C \times P \]

Where: A = Annual soil erosion per unit area (t ha⁻¹), R = Annual Rainfall erosivity EI30 (t m ha⁻¹), K = Soil erodibility (t ha⁻¹ EI30⁻¹), S = Slope steepness, L = Slope Length, C = Cover management factor and P = Conservation measures

In the GBR Source framework, a daily time-step, spatially variable RUSLE was used to generate the hillslope sediment loads for grazing landuse, although if the user wishes it can also applied to other landuses for example cropping and horticulture. Sediment loads are lumped and reported within each Functional Unit (land use) (typically 1 km² – 30 km²) for a given sub-catchment at a daily timestep. However the modelled daily loads for landuses where RUSLE has been applied are derived at a finer scale, with calculations performed on a typically, 30 m by 30 m - 900 m² grid. The pre-processing approach is used to limit model runtime, while still preserving the finer spatial detail. The hillslope erosion model is described below and full details of the method can be found in Ellis (2017).

Rainfall erosivity factor (R) values were calculated using the standard methods outlined in Waters et al. (2014). Catchment daily rainfall used in the hydrology modelling provides the daily rainfall input.

Soil erodibility factor (K) raster was calculated using methods of Loch and Rosewell (1992). Soil data for these calculations was sourced from the Queensland soils database using the best available soils mapping for spatial extrapolation.

Length and Slope factors (LS) L factor was set to 1 for grazing areas and is only applicable where rill erosion can occur. The assumption was that rill erosion is generally not found in low intensity grazing systems. The S-
factor was generated of the 30m shuttle digital elevation model (DEM), using standard methods outlined in Waters et al. (2014).

Cover factor (Cfactor), was supplied at a seasonal timestep (four scenes per year) for the model run period via the remote sensing Fractional Cover Index product (FCI) at a 25 m grid square scale, and resampled to 30m to match the DEM. In addition, the FCI is converted from remotely sensed values to a cover metric that better matches traditional manually observed cover (Trevithick and Scarth 2013) and for grazing lands this is then converted to Cfactors using the methods of Rosewell (1993). However the model provides the option for the use of static Cfactors, or other calculation methods.

Practice (P) factor was treated as a static value and not applied in the model.

The RUSLE model generates erosion at the grid scale, however the model requires the eroded material be split into fine and coarse sediment and delivered to the stream. To do this, loads are multiplied by a clay + silt grid and finally by a sediment delivery ratio (SDR). Although again these options are flexible. The final hillslope erosion equation can be represented in following form.

Total suspended sediment load (kg/day) = RUSLE sediment load (kg/day) * (clay proportion + silt proportion) * SDR

2.3. Model Disaggregation and Evaluation

As the individual grid calculations are not recorded at model runtime, the grid cell derived loads for the study catchment comparisons needed to be calculated externally, allowing appropriate evaluations to be made between Source Catchments modelled loads and the much smaller catchment study data. To perform the spatial calculations, subsets of the relevant grids need to be defined. The Virginia Park and Brigalow Catchment Study monitoring site boundaries were sourced from relevant project staff.

The spatial values for KLS, ground cover and Cfactor, were obtained for each monitored catchment using the python library raster stats (http://pythonhosted.org/rasterstats). For the Brigalow Catchment Study cropping catchment, we used a temporally static Cfactor value of 0.1, obtained from prior modelling of cropping in the Fitzroy (Dougall et al. 2006).

The Rfactor and runoff were extracted from the relevant GBR Source model (McCloskey 2017). The final RUSLE calculations were then performed within a spreadsheet. To test the methodology the externally calculated values for each grid cell were assessed against the aggregated landuse modelled outputs. The calculation showed very minor differences (<0.1%).

Source modelled SDRs are commonly 10% to represent depositional and transport processes at the larger subcatchment scale (McCloskey 2017). SDRs generally decrease with increasing catchment size (Walling 1983). Therefore due to the small catchment size, delivery ratio should be adjusted for better representation, in this instance we arbitrarily doubled the delivery ratios. However it is acknowledged that this may not be suitable for the other catchment studies and may be site specific. Sediment loads and cover were then compared at the average annual and water year scales.

3. RESULTS AND DISCUSSIONS

3.1. Model Performance – Virginia Park

Ground cover and erosion rate comparisons are shown for Virginia Park (Figure 1. a,b,c) (all grazing landuse). At all three flumes measured ground cover is shown to align well with the remote sensing product used in the Source model with differences of around (~ 0-4%) at the average annual scale and (~2-10%) at the annual scale. Likewise, erosion is well simulated at flume 1 and 2, with modelled average annual values simulated in the observed range of (~0.05 to 0.15 t/ha). At the annual scale larger differences are observed for the 2004 and 2010 water year. The largest differences in erosion are noted for flume 3 and here erosion is shown to be well underestimated at both the average annual and annual timestep.
The extreme erosion values reported in flume 3 are attributed to a low cover area immediately above the flume (Bartley et al. 2014). To confirm the low cover values, we interrogated the spatial arrangement of cover in flume 3 (Figure 1. d) which indicates the model has a reasonable spatial depiction, apart from a small area at the bottom of catchment. Within Source, this is represented by two pixels. The modelled value of cover is ~30% while the monitored one is 0%. Importantly, if the model was correctly populated with the monitored cover, the erosion rates would be significantly higher (~1-2 t/ha). However this would still be a likely underestimate, suggesting that the other factors are not being correctly populated in the model. For example, soil erodibility or slope, and we note lower measured slopes than those derived from the Shuttle DEM in the Source model. New remote sensing products show the potential to better identify slope and permanently bare areas, and one could theoretically better populate the model as the data comes online.

At the larger Virginia Park (14 km²) catchment scale site, and the Burdekin catchment, the contribution of hillslope erosion to sediment loads has been assessed using sediment tracing technology (Wilkinson et al. 2013). The sediment tracing work suggests that (~20%) of the load is derived from surface erosion sources. By comparison, modelled hillslope loads comprise approximately ~15% of the observed Virginia Park catchment scale loads, using a SDR value of 10%.

There are many assumptions and limitations involved in small scale comparisons. For instance, side by side replicate plots have monitored differences of ±30% (Risse et al. 1993). Nonetheless on the steeper Virginia Park hillslopes the model appears to be showing some close approximations of hillslope erosion. A potential area for improvement would be better simulation of the permanently bare areas and slope. We note that data at finer scales will become increasingly more available and its impact on hillslope modelling could potentially be tested at the Virginia Park site.

3.2. Model Performance – Brigalow Catchment Study

Erosion rate comparisons for the Brigalow Catchment Study are shown in Error! Reference source not found.. a,b. At the average annual scale (1986-2007), erosion is well simulated for the three landuses (using a SDR of 20%), recorded values are within 3% of the predicted. More importantly and regardless of the SDR the relative differences between landuses were well simulated at this small scale (refer Figure 2a).

At the annual scale (Error! Reference source not found. b), large variation can be observed. Potentially the data suggest that the model may underestimate the larger events, and overestimate the smaller ones which is a common occurrence with this modelling approach (Kinnell 2010). However more detailed investigation is required to identify if rainfall and runoff differences are driving the observed variation. As modelled rainfall comes from a distant gauge (>10km away), investigation of temporal model performance characteristics may

Figure 1. Virginia Park, hillslope flume ground cover and erosion rate comparisons (a) Flume 1, (b) Flume 2, (c) Flume 3 and (d) Flume 3 spatial cover generated from calibrated on ground study imagery and that used in the source model.
be better served by using the rainfall measured at the study site. However rainfall error is an inherent part of lumped catchment modelling in a sparsely monitored environment.

Interestingly, cover measurements are available for the “open grazing” catchment from the period (1987-1990). After this period ground cover measurements ceased due to cover being constantly in the range of (95% - 100%) as a result of good ground cover maintained or A class management practice being implemented. The cover measurements used in the Source model were slightly lower resulting in higher erosion.

Currently, the GBR Source models are populated with outputs from a paddock scale cropping model. The cropping models are difficult to evaluate at larger spatial scales due to a lack of measured data. If the GBR Source application of RUSLE displays similar levels of predictive performance at other well monitored cropping sites, there is potential benefit in using the RUSLE model to generate erosion values for these areas as well. These in turn could be used to test the performance of the paddock scale cropping model across GBR cropping lands.

4. CONCLUSIONS AND RECOMMENDATIONS

Small catchment studies often contain a wealth of information to inform catchment models, including the impact of management practice on constituent loss. However utilising this data to assess and inform large catchment scale models is challenging. The disaggregation and assessment of modelled loads outlined in this paper, facilitates small to large scale hillslope erosion comparisons. For the two case study sites, erosion and cover are shown to be well predicted at the average annual scale and areas for further improvement and investigation have been highlighted. We recommend further sites across the GBR should be assessed. Validating disaggregated modelled data against plot scale experimental data, provides an alternative data set for validation previously not utilised. The approach shows a lot of potential to improve confidence in modelled sediment generation predictions.

REFERENCES


Dougall and McCloskey, Great Barrier Reef Source Catchment’s modelling: Assessing hillslope erosion modelling performance at paddock scale experimental sites


