Assimilating catchment processes with monitoring data to estimate sediment loads to the Great Barrier Reef

D. E. Pagendam\textsuperscript{a*}, P. M. Kuhnert\textsuperscript{b}, W. B. Leeds\textsuperscript{c}, C. K. Wikle\textsuperscript{d}, R. Bartley\textsuperscript{e} and E. E. Peterson\textsuperscript{b}

Quantifying riverine sediment loads, and the uncertainty around these estimates, is important for monitoring the impact of land use on ecologically sensitive receiving waters such as the Great Barrier Reef lagoon. We used a Bayesian Hierarchical Modelling approach that assimilates information from a process model for runoff, a mechanistically motivated statistical model for sediment generation and observed runoff and sediment load data. The approach was trialled on a 10-year dataset collected from a 14-km\textsuperscript{2} sub-catchment in the Burdekin basin, Australia. Using our model, we were able to estimate daily sediment concentrations, discharges and loads (with credible intervals) over a 10-year period, including 3 years where there were long periods of missing observational data. We found that for the high-frequency monitoring undertaken at the study site, credible intervals around sediment loads were narrow. Credible intervals were substantially wider in years where observational data were not available and load estimates relied on the underlying processes and neighbouring observations. The method presented here is the first attempt at assimilating discharge and concentration measurements with process models for the purpose of sediment load estimation. The potential for quantifying loads entering the Great Barrier Reef lagoon is promising, particularly for ephemeral streams that are typical of arid and semi-arid Australia. Copyright © 2014 John Wiley & Sons, Ltd.

**Keywords:** coral reef; data assimilation; model–data fusion; uncertainty quantification; water quality

1. **INTRODUCTION**

The volumes of sediment discharged from Great Barrier Reef (GBR) catchments have increased since European settlement as a consequence of increased soil erosion following agricultural expansion (McCulloch et al., 2003; Kroon et al., 2012; Bartley et al., 2014). Recent scientific studies have shown that these increases in sediment have altered the coral and seagrass communities across the GBR that support dugong, turtle and fish populations (De’ath et al., 2012; Fabricius, 2005). This is of primary interest to managers and a focus of state and national initiatives such as the Reef Water Quality Protection Plan, where a 20% reduction in sediment is required by 2020 (Reef Water Quality Protection Plan Secretariat, 2013).

Reliably quantifying the sediment loads generated from GBR catchments is an important first step towards identifying (i) the size of the load, (ii) the major sources of erosion contributing to the load and (iii) providing some confidence in the management strategies implemented to mitigate erosion impacts on the GBR lagoon (e.g. (Bartley et al., 2010b)).

Sediment loads represent the mass of sediment transported past a particular site on a river over a specific period. Mathematically, a sediment load (in tonnes) is computed as the integral of the concentration, \( c(t) \), (in tonnes per cubic metre of water) and flow, \( q(t) \), (in cubic metres of water per second) over the time interval \([t_1, t_2]\) such that

\[
L = \int_{t_1}^{t_2} c(t)q(t)\,dt.
\]

In practice, measurements are discrete, and this integral is approximated as a sum. When data are on a daily time step, the load is calculated as

\[
L = \sum_{t=t_1}^{t_2} \mu_c(t)\,v(t),
\]

where \( \mu_c(t) \) represents the average sediment concentration on day \( t \), and \( v(t) \) is the total discharge volume on day \( t \).

Loads cannot be measured directly and are estimated using derived flow and total suspended sediment (TSS) concentration data. While the installation of monitoring equipment provides the best chance of reliably estimating a load, even state-of-the-art monitoring equipment...
is subject to measurement error. Furthermore, when collecting hydrographic data, it is common for equipment to malfunction, be subject to vandalism or result in damage during large runoff events and storms. Such phenomena can lead to gaps in the time series collected at monitoring stations and highlights the need for methodologies for interpolating between measurements to ensure a load can be estimated.

Until now, the quantification of loads to the GBR lagoon has either been empirically based, (Mitchell et al., 2006; Bainbridge et al., 2007), stochastic, whereby estimation is achieved through a mechanistically motivated statistical model (Wang et al., 2011; Kuhnert et al., 2012) or entirely deterministic, without representation of the uncertainty in load estimates (McKergow et al., 2005; Kinsey-Henderson et al., 2007; Armour et al., 2009; Wilkinson et al., 2014). Bayesian Hierarchical Modelling (BHM) provides a framework through which data can be assimilated with process-based models in a statistically rigorous way. The approach falls into a broad class of modelling techniques often branded as data assimilation, physical-statistical modelling, data fusion or model–data fusion, but where the underlying goal is to combine information from multiple data, model and information sources in order to make inferences in a unifying statistical framework. The methods we employ herein are based on a BHM framework (Berliner, 1996; Wikle and Berliner, 2007; Cressie and Wikle, 2011) that allows TSS concentrations, discharges and sediment loads to be estimated while accounting for errors in (i) observed data, (ii) physical processes underpinning the measurements and (iii) model parameters. There have been numerous articles that focus on uncertainty assessment of stream discharges from rainfall-runoff models (Wu et al., 2010; Vrugt et al., 2008; Kuczera et al., 2006; Beven and Freer, 2001), uncertainty of erosion processes (Falk et al., 2009) and uncertainty assessment of models for load generation (Doherty and Johnston, 2003; McCloskey et al., 2011; Ellis et al., 2009). Some of these adopt a Monte Carlo simulation approach, where parameters of a model are simulated according to a generalised likelihood (Beven and Freer, 2001; Doherty and Johnston, 2003; McCloskey et al., 2011; Ellis et al., 2009), while others adopt a BHM approach that places priors on parameters of the process that they model and the forcing variables (e.g. rainfall) that drive the processes (Vrugt et al., 2008; Kuczera et al., 2006). In recent North American studies, Bayesian statistical approaches for modelling sediment transport have been developed (Schmelter et al., 2011; Schmelter et al., 2012; Schmelter and Stevens, 2013), allowing uncertainties in model parameters, data and model predictions to be rigorously quantified. However, the lack of separate process and data models in those studies means that they do not fit with the BHM framework used herein. Only Wu et al. (2010) present a physical-statistical modelling framework applied to a hydrological model that aligns with our current thinking around the problem.

We present a BHM method for integrating measurements with modelled output to quantify catchment sediment loads with credible intervals. The methods can be applied to (but are not limited to) streams that are ephemeral (i.e. the stream may stop flowing for extended periods), which is a characteristic of many arid and semi-arid rivers around the world. We focus on a single site located on Weany Creek in the Upper Burdekin catchment, where 10 years of flow and TSS data were collected. The latent processes of discharge and concentration were modelled using a process-based model of stream discharge and a mechanistically motivated, statistical model of suspended sediment generation, respectively. The models were validated and model performance evaluated in both the wet and dry seasons. In addition, we explored the measurement errors and compared them with errors estimated from the concentration and discharge processes. Although our primary objective was prediction, we also demonstrate how the model results can be used to learn about parameters from both process models. Finally, we discuss the implications for this modelling framework to be applied to other sites in the GBR catchments to inform management decisions.

2. CASE STUDY: POLLUTANT LOADS ESTIMATION FOR THE WEANY CREEK CATCHMENT

2.1. Study site characteristics

Weany Creek is a 14-km² headwater catchment (S19°53'06.79", E146°32'06.65") within the larger Burdekin basin (approximately 130 000 km²; Figure 1). Weany Creek has a mean channel bed slope of approximately 0.5% and valley hillslopes averaging around 4%. The site was identified as an erosion ‘hotspot’ with respect to impact on the GBR (Prosser et al., 2001), and a research site was established in 1999 to identify erosion sources (Bartley et al., 2007; Kinsey-Henderson et al., 2005; Wilkinson et al., 2013), and responses to grazing land management (Bartley et al., 2010b; Bartley et al., 2010a; Bartley et al., In Press). Data from this site were suitable for this study because of the availability of high-resolution flow, turbidity and sediment concentration records over a 10-year period (2000–2010). Weany Creek is ephemeral and generally only flows between December and April each year. Occasionally, out of season events occur, and therefore, hydrology data are often reported from 1 July to 30 June each year. The 10-year mean annual rainfall for the site was 658 mm (standard deviation of 317 mm). This is slightly higher than the long-term average rainfall (1901–2011) for the nearby Fanning River rain gauge of 604 mm (http://www.nrm.qld.gov.au/silo). During the 10-year study period, the highest rainfall at Weany Creek was 1224 mm in the 2008/2009 wet season, and the lowest rainfall was 303 mm in 2002/2003. The soils in the catchment range from red chromosols on the upper slopes and yellow to brown texture contrast soils with dispersive B-horizons on the lower footslopes. The canopy vegetation is composed primarily of ironbark/bloodwood communities (e.g. Eucalyptus crebra and Corymbia erythrophloia), which are located primarily on the mid and upper slopes. The lower slope sodic soil communities are dominated by more shrubby species (e.g. Carissa ovata and Eremophila mitchelli). The ground cover is dominated by the exotic but naturalised stoloniferous grass Indian Couch (Bothriochloa pertusa). Cattle grazing is the dominant land use in the catchment, although small areas of gold mining have occurred in the upper parts of the catchment over the last century.

2.2. Field instrumentation and data collection

An automatic gauging station was installed at the outlet of Weany Creek in 1999. The first year of data was used to trial the instruments and were not used in this analysis. The station operated successfully for most of 2001–2010, however, a tree destroyed part of the station
in 2007 reducing the reliability of data collected between 2007 and 2008. The station recorded rainfall, stage (i.e. the height of the water), velocity and turbidity (measured in nephelometric turbidity units) at one-minute intervals during events. An ‘event’ was defined when the flow increased beyond about 1 m$^3$ s$^{-1}$. The number of events varied from two events in 2002/2003, to more than 10 events in 2009/2010. To estimate discharge, regressions were fit between velocity and depth, and then the regressed velocity was multiplied by the cross-sectional area at each depth during events. To estimate sediment concentration, a 1-L water sample was collected at programmed intervals across the hydrograph. A linear relationship between total suspended sediment and turbidity was developed using data from 2000–2006 and applied to the 2000-2006 period. A new turbidity sensor was installed in 2007, and a revised relationship was applied from 2007–2012. These relationships were then used to calculate the annual suspended sediment load for each year. In previous publications from this study site, when data were missing (because of station damage), TSS concentrations were linearly interpolated from TSS grab samples. In this paper, we used a mechanistically motivated statistical model to stochastically in-fill missing sediment concentrations. Further details of the monitoring equipment and water sampling design of the gauging site are given in Bartley et al. (2007), Bartley et al. (2010b) and Hawdon et al. (2009). Bedload, which is often considered an important aspect of sediment load estimation, was not measured at the catchment outlet.

3. THE BAYESIAN HIERARCHICAL MODELLING FRAMEWORK

We adopt a Bayesian Hierarchical Modelling framework and follow the concepts of Berliner (1996), Wikle and Berliner (2007) and Cressie and Wikle (2011) to construct a model for quantifying sediment loads at a single river site. Let $Z = (Z_1, \ldots, Z_n)$ represent the random variables corresponding to data obtained in observing the stochastic process $Y = (Y_1, \ldots, Y_n)$, with process model parameters $\theta_p$ and statistical parameters, $\theta_s$. The joint probability density of all components in the model can be conveniently written as the product of three conditional probability density functions, which represent different tiers in the modelling hierarchy

$$ [Z, Y, \theta_m, \theta_s] = [Z|Y, \theta_s] [Y|\theta_m] [\theta_s, \theta_m] $$

(1)

where $[\cdot]$ is shorthand for the probability density function of the enclosed random variables, $[Z|Y, \theta_s]$ is the data model that reconciles the observed data with an underlying latent process, $Y$, or mathematically, the conditional density of the observations given the latent process and statistical parameters; $[Y|\theta_m]$ is the process model or the conditional density of the latent process given the model parameters; and $[\theta_s, \theta_m]$ is the parameter model or prior probability distribution over the model and statistical parameters. For ease of reference, the three tiers of models outlined in the succeeding texts are summarised in Table 1.

3.1. Data model

Let $Z^Q = (Z_1^Q, \ldots, Z_n^Q)$ and $Z^{TSS} = (Z_1^{TSS}, \ldots, Z_n^{TSS})$ represent random vectors corresponding to observed, daily time series of stream discharge volumes (in m$^3$) and TSS (mg L$^{-1}$), respectively. Furthermore, let $z^Q = (z_1^Q, \ldots, z_n^Q)$ and $z^{TSS} = (z_1^{TSS}, \ldots, z_n^{TSS})$ represent observed realisations of $Z^Q$ and $Z^{TSS}$, respectively; $I_{Y_t^Q}$ be a latent indicator variable that takes the value 1 if the stream is flowing at time $t$ and zero otherwise; and $I_{Y_t^{TSS}}$ be a latent indicator variable that takes the value 1 if there is generation of suspended sediment at time $t$ and the value 0 otherwise. The data model for these quantities can be written as
Table 1. Summary of the three tiers of the Bayesian hierarchical model

<table>
<thead>
<tr>
<th>Tier</th>
<th>Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data models</td>
<td>$Z_t^O</td>
</tr>
<tr>
<td>Process models</td>
<td>$Z_t^{TSS}</td>
</tr>
<tr>
<td>Parameter models</td>
<td>$\theta_1 \sim U(0.0, 5.0)$</td>
</tr>
<tr>
<td></td>
<td>$\theta_2 \sim U(0.0, 5.0)$</td>
</tr>
<tr>
<td></td>
<td>$\theta_3 \sim U(1.0, 500.0)$</td>
</tr>
<tr>
<td></td>
<td>$\theta_4 \sim U(0.0, 400.0)$</td>
</tr>
<tr>
<td></td>
<td>$\theta_5 \sim U(0.0, 10.0)$</td>
</tr>
<tr>
<td></td>
<td>$\theta_6 \sim U(0.0, 1.0)$</td>
</tr>
<tr>
<td></td>
<td>$\theta_7 \sim U(0.0, 1.0)$</td>
</tr>
<tr>
<td></td>
<td>$\theta_8 \sim U(0.0, 1.0)$</td>
</tr>
<tr>
<td></td>
<td>$\theta_9 \sim U(0.0, 1.0)$</td>
</tr>
<tr>
<td></td>
<td>$\beta_i \sim N(0.0, (1000.0)^2)$</td>
</tr>
<tr>
<td></td>
<td>$\phi_i^Q \sim N(0.0, (1000.0)^2)$</td>
</tr>
<tr>
<td></td>
<td>$\phi_i^{TSS} \sim N(0.0, (1000.0)^2)$</td>
</tr>
<tr>
<td></td>
<td>$\gamma \sim U(0.0, 1.0)$</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{Y^O} \sim U(0.0, 10^6)$</td>
</tr>
<tr>
<td></td>
<td>$\sigma_{Y^{TSS}} \sim U(0.0, 5000.0)$</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2_{Z_t^O}$ is time-varying, estimated from data and fixed</td>
</tr>
<tr>
<td></td>
<td>$\sigma^2_{Z_t^{TSS}}$ is time-varying, estimated from data and fixed</td>
</tr>
</tbody>
</table>

\[
Z_t^O | Y_t^O, y_t^O, \sigma^2_{Z_t^O} \sim \begin{cases} 
TN_0^\infty \left( h \left( \max \{y_t^O, \epsilon_1, \sigma^2_{Z_t^O}, \sigma^2_{Z_t^O}\} \right) , \sigma^2_{Z_t^O} \right) \\
\delta(0)
\end{cases}
\]

(2)

and

\[
Z_t^{TSS} | Y_t^{TSS}, y_t^{TSS}, \sigma^2_{Z_t^{TSS}} \sim \begin{cases} 
TN_0^\infty \left( h \left( \max \{y_t^{TSS}, \epsilon, \sigma^2_{Z_t^{TSS}}, \sigma^2_{Z_t^{TSS}}\} \right) , \sigma^2_{Z_t^{TSS}} \right) \\
\delta(0)
\end{cases}
\]

(3)

where $\delta(\cdot)$ is shorthand for the Dirac delta function, and $TN_0^b(\mu, \sigma^2)$ denotes a normal distribution with mean $\mu$ and variance $\sigma^2$, truncated on the left at $a$, truncated on the right at $b$ and having probability density function:

\[
f_X \left( x; \mu, \sigma^2 \right)^b_a = \left( \Phi_X \left( b; \mu, \sigma^2 \right) - \Phi_X \left( a; \mu, \sigma^2 \right) \right)^{-1} \left( 2\pi\sigma^2 \right)^{-\frac{1}{2}} e^{-\frac{(x-a)^2}{2\sigma^2}} \quad (-\infty \leq a < b \leq \infty, x \in [a, b])
\]

In the previous texts, $\Phi_X \left( x; \mu, \sigma^2 \right) := P(X \leq x)$ is the cumulative distribution function of a normally distributed random variable, $X$, with mean $\mu$ and variance $\sigma^2$ and $\Phi_X \left( \infty; \mu, \sigma^2 \right)$ should be interpreted as $\lim_{x \to \infty} \Phi_X \left( x; \mu, \sigma^2 \right)$. The quantities, $y_t^O$ and $y_t^{TSS}$ are...
realisations from the latent processes for stream discharge volume and TSS at time \( t \) (defined in the subsequent section). The parameters \( \sigma_{Z^i_t}^2 \) and \( \sigma_{Z^{TSS}_t}^2 \) represent measurement error variances (assumed to be known and time-varying) for daily discharge volume and TSS, respectively. The function \( h(\cdot) \) is a bias correction function that ensures that the expected value of the resulting random variable is equal to the realisation of the underlying latent process, unless \( y^*_t = 0 \), in which case, the expected value is corrected to be an arbitrarily small value, \( \epsilon = 10^{-6} \) (herein). The need for \( \epsilon \) arises because for the \( TN_0^\infty \) distribution, as \( y^*_t \to 0, h(y^*_t) \to -\infty \). The result of this bias correction is that \( \text{E} \left( Z^i_t \mid y^*_t = 1, \mu^i_t, \sigma^2_{Z^i_t} \right) = \max \{ y^i_t, \epsilon \} \) and \( \text{E} \left( Z^{TSS}_t \mid y^{TSS}_t = 1, \mu^{TSS}_t, \sigma^2_{Z^{TSS}_t} \right) = \max \{ y^{TSS}_t, \epsilon \} \), which admits a small (and negligible) bias for small process values, \( y^*_t \). We borrow the idea of a bias correction approach for the truncated normal distribution from Cangelosi and Hooten (2009). Finally, in the previous model formulation, we make the realistic assumption that it is not possible to make a recording of stream flow when the stream is not flowing.

### 3.2. Process model

Our process models capture the dynamics of stream discharge and sediment concentration using a rainfall-runoff model and a mechanistically motivated statistical model, respectively. The SIMHYD rainfall-runoff model of Chiew et al. (2002) is used to model stream discharge volumes on a daily time step. SIMHYD consists of a system of deterministic equations with nine input parameters, \( \theta = (\theta_1, \theta_2, \ldots, \theta_9) \), outlined in Table 2. Let \( Y^O = (Y^O_1, \ldots, Y^O_n) \) be a random vector from the latent processes of total stream discharge (in m\(^3\)) and \( Y^{TSS} = (Y^{TSS}_1, \ldots, Y^{TSS}_n) \) (which appears in Equation (2)) be a realisation of this process. We write the process model for the total daily stream discharge as a zero-inflated truncated normal distribution using the following specification:

\[
\left[ Y^O_t \mid p^O_t \right] \sim \text{Bernoulli} \left( p^O_t \right)
\]

\[
Y^O_t \mid \mu^O_t, \sigma^2_{Y^O_t} \sim \left\{ \begin{array}{ll}
\text{TN}_0^\infty \left( h \left( \max \{ \mu^O_t, \epsilon \}, \sigma^2_{Y^O_t} \right), \sigma^2_{Y^O_t} \right) & \text{if } Y^O_t = 1 \\
0 & \text{if } Y^O_t = 0
\end{array} \right.
\]

where \( p^O_t \) is the probability of the stream flowing on day \( t \), \( (\mu^O_t, g_t) = f (x_t^*, x^P_t, g_{t-1}, \theta) \), \( f (\cdot) \) is the transition function for the SIMHYD model, \( (\mu^O_t, g_t) \) is a vector containing the predicted total flow and groundwater store at time \( t \) given input variables of rainfall, \( x^P_t \), Morton’s wet environment evapotranspiration, \( x^P_t \), and groundwater store \( g_{t-1} \) at time \( t-1 \). The function \( h(\cdot) \) is the same bias correction function described in the data model and \( \epsilon = 10^{-6} \). Time series of Morton’s wet environment evapotranspiration estimates (Morton, 1983) were obtained from the Data Drill climate database (Jeffrey et al., 2001), whilst time series of rainfall were obtained from a weather station installed at the study site, with occasional gaps in the series filled using the SILO database. Herein, these forcing variables are treated as being error free.

The probabilities of non-zero stream discharge are modelled using a simple logistic function of lagged rainfall

\[
p^O_t := \frac{\exp (-\phi^O_0 - \phi^O r_t)}{1 + \exp (-\phi^O_0 - \phi^O r_t)}
\]

where \( \phi^O = (\phi^O_0, \phi^O) \) is a vector of parameters, \( \phi^O = (\phi^O_1, \ldots, \phi^O_{11}) \), and \( r_t = (x_t^*, x^P_t, g_{t-1}, \ldots, x^P_{t-10}) \) is a vector of lagged rainfall observations.

A mechanistically motivated statistical model is used to represent the sediment generation process of the system. Let \( Y^{TSS}_t = (Y^{TSS}_1, \ldots, Y^{TSS}_n) \) be a random vector from the latent processes generating TSS in the stream (in mg L\(^{-1}\)) and \( Y^{TSS}_t = (Y^{TSS}_1, \ldots, Y^{TSS}_n) \)

---

**Table 2. SIMHYD input parameters**

<table>
<thead>
<tr>
<th>Parameter</th>
<th>Description</th>
<th>Lower bound</th>
<th>Upper bound</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \theta_1 )</td>
<td>Impervious threshold</td>
<td>0.0</td>
<td>5.0</td>
</tr>
<tr>
<td>( \theta_2 )</td>
<td>Rainfall interception store capacity</td>
<td>0.0</td>
<td>5.0</td>
</tr>
<tr>
<td>( \theta_3 )</td>
<td>Soil moisture store capacity</td>
<td>1.0</td>
<td>500.0</td>
</tr>
<tr>
<td>( \theta_4 )</td>
<td>Infiltration coefficient</td>
<td>0.0</td>
<td>400.0</td>
</tr>
<tr>
<td>( \theta_5 )</td>
<td>Infiltration shape</td>
<td>0.0</td>
<td>10.0</td>
</tr>
<tr>
<td>( \theta_6 )</td>
<td>Interflow coefficient</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>( \theta_7 )</td>
<td>Recharge coefficient</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>( \theta_8 )</td>
<td>Baseflow coefficient</td>
<td>0.0</td>
<td>1.0</td>
</tr>
<tr>
<td>( \theta_9 )</td>
<td>Pervious fraction</td>
<td>0.0</td>
<td>1.0</td>
</tr>
</tbody>
</table>

---


wileyonlinelibrary.com/journal/environmetrics
(which appears in Equation (3)) be a realisation of this process. As for stream discharge, the process model for TSS is modelled as a zero-inflated normal distribution, in which we construct in much the same way as before

\[ I_{Y_{TSS}} | p_{t}^{TSS} \sim \text{Bernoulli} \left( p_{t}^{TSS} \right) \]

\[ y_{t}^{TSS} | I_{Y_{TSS}}, \mu_{t}^{TSS}, \sigma_{Y_{TSS}}^{2} \sim \left\{ \begin{array}{ll}
\mathcal{N}_{\infty} \left( h \left( \max \{ \mu_{t}^{TSS}, \epsilon \} , \sigma_{Y_{TSS}}^{2} \right) \right) & (I_{Y_{TSS}} = 1) \\
\delta(0) & (I_{Y_{TSS}} = 0)
\end{array} \right. \]

where

\[ \mu_{t}^{TSS} = \beta_{0}^{TSS} + \beta_{1}^{TSS} \left( s_{t}^{TSS} \right)^{T} + \beta_{5}^{TSS} y_{t}^{Q} + \beta_{6}^{TSS} \left( y_{t}^{Q} \right)^{2} + \beta_{7}^{TSS} v_{t} + \beta_{8}^{TSS} d_{t} \]  

\[ d_{t} := \frac{\sum_{i=1}^{t} Y_{t}^{i-1} x_{t}^{i}}{\sum_{i=1}^{t} Y_{t}^{i-1}} \]  

Terms introduced in Equation (5) are based on important mechanisms in the generation of suspended sediment as identified by Kuhnert et al. (2012) and Wang et al. (2011). In the previous texts, \( s_{t} = (\sin(\omega t), \cos(\omega t), \sin(2\omega t), \cos(2\omega t)) \), with \( \omega := 2\pi / 365.25 \); \( v_{t} \) represents remotely sensed ground cover index (Scarth et al., 2006) in the catchment at time \( t \); \( \beta^{TSS} = \left( \beta_{0}^{TSS}, \beta_{1}^{TSS}, \ldots, \beta_{8}^{TSS} \right) \) and \( \beta_{i}^{TSS} = (\beta_{i1}^{TSS}, \ldots, \beta_{i3}^{TSS}) \), which makes the Markov chain Monte Carlo algorithm computationally burdensome and (ii) as rain drives flow, rain could also be considered a driver of sediment generation. The probabilities of non-zero TSS are again modelled using a simple logistic function of lagged rainfall

\[ p_{t}^{TSS} := \frac{\exp \left( -\phi_{0}^{TSS} - \phi^{TSS} r_{t} \right)}{1 + \exp \left( -\phi_{0}^{TSS} - \phi^{TSS} r_{t} \right)} \]

where \( \phi^{TSS} = (\phi_{0}^{TSS}, \phi^{TSS}) \) is a vector of parameters with \( \phi^{TSS} = (\phi_{11}^{TSS}, \ldots, \phi_{13}^{TSS}) \), and \( r_{t} \) is a vector of lagged rainfall observations as in Equation (4).

3.3. Parameter model

The list of parameters in this model and their prior distributions are listed in Table 1. Most parameters in the model were assigned weakly informative priors. We assigned informative, uniform priors to the process model standard deviations \( \sigma_{Y_{Q}} \) and \( \sigma_{Y_{TSS}} \) to ensure that these were constrained to a physically realistic range. The time-varying measurement error variances on the observed TSS, \( \sigma_{Z_{1}^{TSS}}^{2} \) and stream discharge volume, \( \sigma_{Z_{2}^{Q}}^{2} \), were assumed known for the purpose of model identifiability, and their values were determined using the methods outlined in the succeeding texts.

3.3.1. Prior for errors in flow

Monitoring equipment installed at the gauging station on Weany Creek recorded stage and stream velocity at 1-minute intervals during events. Rating curves were developed for the gauging station through two relationships: (i) a stage-velocity curve and (ii) a stage-area curve. The stage-area relationship consisted of the following relationship:

\[ a_{i} = \beta_{1}^{a} h_{i} + \beta_{2}^{a} h_{i}^{2} + \beta_{3}^{a} h_{i}^{3} + \beta_{4}^{a} h_{i}^{4} \quad (i = 1, \ldots, n_{R}; h_{i} > 0) \]

with \( \beta_{a} = (\beta_{1}^{a}, \ldots, \beta_{4}^{a}) \) as parameters requiring estimation and \( (h_{i}, a_{i}) \) as paired observations of stage and the stream wetted cross-sectional area taken at some juncture. The polynomial model in Equation (7) was a near perfect fit to the observed data and therefore a deterministic model was deemed appropriate. A stochastic model was required, however, for the relationship between velocity and stage

\[ \log(V_{i})|h_{i} \sim \mathcal{N} \left( \mu_{V_{i}}, \sigma_{V}^{2} \right) \]

where

\[ \mu_{V_{i}} = \beta_{0}^{V} + \beta_{1}^{V} \log(h_{i}) + \beta_{2}^{V} \log(h_{i})^{2} + \beta_{3}^{V} \log(h_{i})^{3} + \beta_{4}^{V} \log(h_{i})^{4} + \beta_{5}^{V} \log(h_{i})^{5} \]
The quantities $\beta^V = (\beta^V_0, \ldots, \beta^V_K)$ and $\sigma^V$ required estimation. The random variable for velocity given stage ($V_i | h_i$), therefore followed a log-normal distribution and the random variable for instantaneous stream flow (in m$^3$ s$^{-1}$) given stage ($Q_i | h_i$) was simply

$$Q_i | h_i := a_i \times V_i | h_i$$

These relationships are exhibited in Figure 2.

Using these relationships, we calculated the measurement error variance for each individual measurement recorded at the monitoring site for a given day and used these to obtain the error variance on the total daily flow as the sum of independent random variables. More formally, let $\hat{\mu}_{q,t,1}, \hat{\mu}_{q,t,2}, \ldots, \hat{\mu}_{q,t,m}$ be the $m$ expected stream velocities (m s$^{-1}$) for day $t$ given observed stages $h_{t,1}, h_{t,2}, \ldots, h_{t,m}$. Furthermore, let $\hat{\sigma}^2_{q,t,1}, \hat{\sigma}^2_{q,t,2}, \ldots, \hat{\sigma}^2_{q,t,m}$ be the measurement error variances associated with these estimates of velocity. From the regression model outlined previously, we compute $\hat{\mu}_{q,t,i} = e^{\mu V_i + \sigma^2 V_i/2}$ and $\hat{\sigma}^2_{q,t,i} = (e^{2\mu V_i} - 1) e^{2\mu V_i + \sigma^2 V_i}$ as the mean and variance of a log-normally distributed random variate on the natural scale. We computed the daily observed discharge volume as $Z_t = \sum_{i=1}^{m} \Delta t a_i \hat{\mu}_{q,t,1}$ and $\sigma^2_{Z_t} = \text{Var}(Z_t | Y_t^Q) = \sum_{i=1}^{m} \Delta t^2 a_i^2 \hat{\sigma}^2_{q,t,i}$. The quantity $\Delta t$ is the time interval (in seconds) between the $i$th observation and the next, and $a_i$ is the wetted cross-sectional area of the stream (in m$^2$) for stage $h_{t,i}$. The coefficients of variation for these measurement errors are shown in Figure 3(a).

### 3.3.2. Prior for errors in total suspended sediment

A TSS–turbidity relationship was used to obtain estimates of mean daily TSS and associated error variances. Two linear relationships, for two different periods, were delineated from this data. The first relationship was based on data collected from 2000 to 2006. The second relationship was based on data collected between 2009 and 2011. Data from 2007 to 2009 were deemed to be unreliable following a tree falling on the turbidity sensor. As there was no observed turbidity data for this period, no TSS–turbidity relationship was needed between these dates.

The TSS–turbidity relationships were applied to turbidity measured at 1-minute intervals at Weany Creek to provide daily estimates of TSS. Let $(U_i, C_i)$ denote the random variables corresponding to paired observations of turbidity and TSS at some juncture and $(u_i, c_i)$, their respective observations. We model the TSS-turbidity relationship as

---

**Figure 2.** Stage-area (a) and stage-velocity (b) relationships used to estimate daily stream discharge and associated measurement errors. Dashed lines show the fitted model

**Figure 3.** Relationships between observations and the measurement-error coefficient of variation for (a) daily discharge volumes and (b) mean daily total suspended sediment
The forms of the conditional distributions required for the MH sampling are specified as follows:

\[ C_i \sim N \left( \mu_{C_i}, I_i \sigma_{C,1}^2 + (1 - I_i) \sigma_{C,2}^2 \right) \]

where

\[ \mu_{C_i} = I_i \left( \beta_{C,1}^c \right) + (1 - I_i) \left( \beta_{C,2}^c \right) \]

\( I_i \) is an indicator function that takes that value 1 if the tuple \((U_i, C_i)\) was observed between 2000 and 2006 and zero otherwise; \( \beta_{C,1}^c, \beta_{C,2}^c, \beta_{C,1}^c, \beta_{C,2}^c \) and \( \sigma_{C,1}^2, \sigma_{C,2}^2 \) were variance parameters requiring estimation from the data. These two relationships are exhibited in Figure 4.

Let \( \mu_{c,t,1}, \mu_{c,t,2}, \ldots, \mu_{c,t,m} \) be the \( m \) estimated TSS (mg L\(^{-1}\)) for day \( t \) given observed turbidities \( u_{t,1}, u_{t,2}, \ldots, u_{t,m} \). Furthermore, let \( \sigma_{c,t,1}^2, \sigma_{c,t,2}^2, \ldots, \sigma_{c,t,m}^2 \) be the measurement error variances associated with these estimates of TSS. Once again, assuming independence of the individual measurements, the mean daily TSS volume was then computed as the time-weighted average \( Z_{i,\text{TSS}} = \sum_{i=1}^{m} \Delta_i \mu_{c,t,1} / \sum_{i=1}^{m} \Delta_i \) and the corresponding time-weighted standard error as \( \sigma_{Z_{i,\text{TSS}}}^2 = \text{Var} \left( Z_{i,\text{TSS}} \mid Y_{i,\text{TSS}} \right) = \sum_{i=1}^{m} \Delta_i \sigma_{c,t,1}^2 / (\sum_{i=1}^{m} \Delta_i)^2 \). The coefficients of variation for these measurement errors are shown in Figure 3(b).

### 3.4. Estimation

Joint estimation of the model parameters, \( \theta_m = \left( \beta_{\text{TSS}}, \phi_{\text{TSS}}, \phi_{\text{Q}}, \gamma, \sigma_{\text{TSS}}^2, \sigma_{\text{Q}}^2 \right) \) and the latent state variables \( Y_{\text{Q}} \) and \( Y_{\text{TSS}} \) was achieved by sampling from the posterior distribution \( \left[ Y_{\text{Q}}, Y_{\text{TSS}}, \beta_{\text{TSS}}, \phi_{\text{TSS}}, \phi_{\text{Q}}, \gamma, \sigma_{\text{TSS}}^2, \sigma_{\text{Q}}^2 \right] \) via a Metropolis–Hastings (MH) within Gibbs sampling scheme (Metropolis et al., 1953; Hastings, 1970; Tierney, 1994). The approach was implemented in R (R Core Team, 2012) but with individual MH samplers coded in C++ using the Rcpp package (Eddelbuettel and François, 2011; Eddelbuettel, 2013). The forms of the conditional distributions required for the MH sampling are specified as follows:

\[
\begin{align*}
Y_{t,\text{TSS}}, Y_{t,\text{Q}} & \propto Y_{t,\text{TSS}} \left[ Y_{t,\text{TSS}}, \beta, \phi_{\text{TSS}}, \gamma, \sigma_{\text{TSS}}^2 \right] \\
Y_{t,\text{Q}} & \propto Y_{t,\text{TSS}} \left[ Y_{t,\text{TSS}}, \beta, \phi_{\text{TSS}}, \gamma, \sigma_{\text{TSS}}^2 \right] \\
|\theta| & \propto |\theta| \prod_{t=1}^{n} Y_{t,\text{Q}}, \beta \phi_{\text{TSS}}, \gamma, \sigma_{\text{TSS}}^2 \\
|\beta \phi_{\text{TSS}}| & \propto |\beta \phi_{\text{TSS}}| \prod_{t=1}^{n} Y_{t,\text{TSS}}, Y_{t,\text{Q}}, \beta \phi_{\text{TSS}}, \gamma, \sigma_{\text{TSS}}^2 \\
|\phi_{\text{TSS}}| & \propto |\phi_{\text{TSS}}| \prod_{t=1}^{n} Y_{t,\text{TSS}}, Y_{t,\text{Q}}, \beta \phi_{\text{TSS}}, \gamma, \sigma_{\text{TSS}}^2 \\
|\gamma| & \propto |\gamma| \prod_{t=1}^{n} Y_{t,\text{TSS}}, Y_{t,\text{Q}}, \beta \phi_{\text{TSS}}, \gamma, \sigma_{\text{TSS}}^2 \\
|\sigma_{\text{TSS}}^2| & \propto |\sigma_{\text{TSS}}^2| \prod_{t=1}^{n} Y_{t,\text{TSS}}, Y_{t,\text{Q}}, \beta \phi_{\text{TSS}}, \gamma, \sigma_{\text{TSS}}^2 \\
|\sigma_{\text{Q}}^2| & \propto |\sigma_{\text{Q}}^2| \prod_{t=1}^{n} Y_{t,\text{TSS}}, Y_{t,\text{Q}}, \beta \phi_{\text{TSS}}, \gamma, \sigma_{\text{TSS}}^2
\end{align*}
\]
\[ [\phi^O | \theta] \propto [\phi^O] \prod_{t=1}^{n} [Y_t^{O}, \tilde{Y}_t^{O} | \theta, \phi^O, \sigma^2_Y^O, x_P^{1:t}, x_R^{1:t}] \]

\[ [\gamma | \theta] \propto [\gamma] \prod_{t=1}^{n} [Y_t^{TSS}, \tilde{Y}_t^{TSS} | Y_t^{O}, \beta^{TSS}, \phi^{TSS}, x_R^{1:t}, \gamma, \sigma^2_Y^{TSS}] \]

\[ [\sigma^2_Y^O | \theta] \propto [\sigma^2_Y^O] \prod_{t=1}^{n} [Y_t^{O}, \tilde{Y}_t^{O} | \theta, \phi^O, r_1, \sigma^2_Y^O, x_P^{1:t}, x_R^{1:t}] \]

and

\[ [\sigma^2_Y^{TSS} | \theta] \propto [\sigma^2_Y^{TSS}] \prod_{t=1}^{n} [Y_t^{TSS}, \tilde{Y}_t^{TSS} | Y_t^{O}, \beta^{TSS}, \phi^{TSS}, x_R^{1:t}, \gamma, \sigma^2_Y^{TSS}] \]

Note that the block update of the pairs \((Y_t^{O}, \tilde{Y}_t^{O})\) and \((Y_t^{TSS}, \tilde{Y}_t^{TSS})\) arises from the need to jointly decide whether the process takes a non-zero value and then what value that should be (if necessary). The quantities \(d_t, p_t^{TSS}\) and \(p_t^{O}\) are not conditioned upon in the previous posterior representations because they are deterministic functions of other random quantities. The dependence of \(Y_t^{TSS}\) on \(x_R^{1:t}\), for example, arises from the mean model for TSS having dependence on \(d_t\) (Equation (5)). Similarly, the probabilities \(p_t^{O}\) and \(p_t^{TSS}\) are computed directly from \(\phi^O\) and \(\phi^{TSS}\), respectively.

All inferences were made from a Markov chain that used the following sequence of phases: (i) an initial 10 000 adaptive burn-in steps during which proposal distributions were tuned to achieve an acceptance rate of 15–30% for all sampled random variables; (ii) a subsequent 40 000 burn-in steps with fixed proposal distributions; and (iii) a final 100 000 iterations with fixed proposal distributions. These final 100 000 iterations were thinned, by taking only every 10th sample in order to reduce the volume of data being stored to disk. Inferences were based on these final 10 000 iterations.

### 4. APPLICATION AT WEANY CREEK

The physical-statistical modelling approach was applied to Weany Creek, where daily measurements of concentration and flow were incorporated into the model along with the measurement errors derived in Section 3. Outputs from the model were in the form of posterior probability distributions, which were obtained for the parameters, predictions (daily sediment concentrations, daily discharges and daily loads) and estimates of annual sediment loads. Of particular interest are the 2007/2008 and 2008/2009 wet seasons where instrumentation failure reduced the reliability of the high-frequency turbidity data.

#### 4.1. Parameter and sediment load estimation

The changes between prior and posterior distribution for all free parameters in the model are shown in Figures 5, 6 and 7. We use kernel density estimates to depict the densities for the posterior distributions in these figures. For all parameters, we observe some degree of ‘learning’ about the credible ranges of values. For some parameters, like SIMHYD parameters \(\theta_1, \theta_2, \theta_3, \theta_4\) and \(\theta_5\), the range of values spanned by the posterior distribution is narrow compared with the prior, indicating strong learning. For others, such as SIMHYD parameter \(\theta_2\), the data have not influenced the posterior greatly, and our learning is much weaker.

The parameters governing the SIMHYD rainfall-runoff model were presented in Figure 5 and may not agree entirely with what a catchment modeller, setting fixed parameter values, might choose. For example, the pervious fraction parameter, \(\theta_9\), has a fairly tight posterior distribution around 0.2. In conceptual terms, this suggests that around 80% of the catchment area behaves as if it were a concreted surface and unable to be penetrated by rainfall, leading to runoff. In calibrating SIMHYD, a catchment modeller might easily have set this parameter equal to 1.0, but our results suggest this would likely be suboptimal and that in actuality, despite having no urban development, the catchment behaves as if a large fraction was impermeable. There are several reasons why the pervious fraction parameter was so high in this model. Firstly, semi-arid rangeland areas can experience runoff coefficients close to 100% particularly during high-intensity events (e.g. Alchin (1983)) and when ground cover is low (e.g. Descheemaeker et al. (2006)). Runoff coefficients as high as 58% have been recorded on hillslopes in this catchment (Bartley et al., 2010a). Secondly, the distribution of rainfall within the catchment can vary considerably, particularly during storm season. If the rainfall is not captured adequately by the rain gauge network, then there is likely to be some error in the runoff coefficient estimate within these catchments. Regardless of the reason, this finding is a reminder that whilst rainfall-runoff models such as SIMHYD are based around a physical process, they are only a coarse representations of reality, and we should be mindful that the parameters in the model do not correspond to real physical quantities.

The process model for concentration included terms that mimicked some of the key hydrological processes as outlined in Wang et al. (2011) and Kuhnert et al. (2012). One such term was the discounting term, \(\gamma\), applied to rainfall. The estimate of \(\gamma\) was quite low (median of 0.283; Figure 6), amounting to a rapid discounting in the influence of past rainfall on TSS at any given time. This is perhaps not unexpected given the ephemeral nature of the stream and markedly different behaviour to that observed in higher order streams as studied in Wang et al. (2011) and Kuhnert et al. (2012).
Cover, a key driver in sediment movement, was a predictor in the model for sediment concentration and was identified as a significant driver of sediment movement in this model. We see from Figure 5 that the probability distribution over the coefficient for cover, $\beta_8$, was strongly negative and did not span zero. This indicated a strong relationship between low cover and higher TSS in the model. This is in agreement with the findings discussed in Bartley et al. (2010b), relating lowest cover with the period immediately before a drought-breaking wet season and therefore, higher than average erosion rates. We also found that whilst $\beta_6$, the parameter controlling the effect of discharge on TSS, tended to be positive, the opposite was true of $\beta_7$, which controlled the effect of squared discharge on TSS. This suggested that for larger discharge events, the TSS was sub-linear with flow, indicating a dilution of suspended sediment at higher discharge volumes.

To accurately model the stream dynamics at Weany Creek, it was necessary to estimate the probability of non-zero discharge and non-zero TSS using rainfall data. Posterior samples of these probabilities are presented in Figure 8. It is clear that the probabilities are almost identical for discharge and TSS. Posterior samples of discharge and TSS series (along with the observational data) are presented in Figures 9 and 10. We draw the readers’ attention to the data gaps in the years 2007, 2008 and 2009 and the black trajectories of TSS that have been predicted during these periods. The trajectories exhibit substantial variability in these gaps, acknowledging uncertainties inherent in the process dynamics for TSS. Close examination of these figures revealed that the trajectories follow the data closely but not perfectly, acknowledging the observation errors associated with these observations. Whilst no data gaps existed in the discharge data, close examination of the
Figure 7. Prior (red) and posterior (black) probability density functions for the parameters used to estimate $p_t^Q$ and $p_t^{TSS}$.

Figure 8. Plots of (a) daily rainfall, (b) posterior samples of the probability series $p_t^Q$ and (c) posterior samples of the probability series $p_t^{TSS}$.

Trajectories revealed realistic variability in discharge dynamics. The estimated probabilities of non-zero stream discharge and concentration for Weany Creek appeared to closely reflect the ephemeral behaviour exhibited in the observed data. Whilst we have included a model for zero discharge and concentration, this part of the model could easily be omitted from the process model to make the methods suitable for perennial streams found further down the stream network and in wetter climates.
The Weany Creek site is an intensively sampled site. The volume of data (particularly TSS) that is available at this site is far greater than for many other sites. Figure 11 shows the estimated sediment loads by calendar years 2001–2010 and by financial years (1 July 1 year to 30 June the following year) within the same 10-year period. We report loads using both calendar and financial years to allow for direct comparison with different reporting practices used in previous studies. Estimates were computed from the predictions of discharge and TSS from the posterior distribution (i.e. our predictions based on the priors, process model and observed data). We see good agreement between the loads computed from observational data alone (i.e. the sum of the product of observed daily discharges and observed sediment concentrations) and the medians of the posterior distributions. The uncertainties that we have estimated in this model could be considered an example of what to expect under optimal monitoring conditions (i.e. when you have high-frequency monitoring). The loads predicted through this approach were higher than those reported in previous studies (Bartley et al., In Press), however, we believe that previous studies underestimated some loads through the interpolation methods employed when monitoring equipment failed. Encouragingly, the posterior distributions for the loads...
produced in the present study tend to cover previous load estimates at this site. As we move towards sites that are monitored infrequently through time, we would expect to see larger uncertainties in the estimated loads. For example, during the 2007/2008 and 2008/2009 wet seasons when monitoring data were sparse, we saw marked increases in the size of the credible intervals for the loads compared with other years. For sites that are data limited, larger uncertainties around load estimates should be expected. This approach highlights the benefit of high-frequency data for reliably quantifying sediment loads (and uncertainties). Given that we are now able to rigorously quantify these errors, managers can begin to use this information to make informed decisions about whether to invest in further monitoring activities.

4.2. Validation study

As part of the analysis, we also performed a validation study where observational data were omitted from the analysis for a dry calendar year (2001) and a wet calendar year (2010). For each of these years, we firstly omitted all TSS observations, and secondly, all TSS and discharge
Figure 11. Posterior load estimates over (a) 10 calendar years and (b) nine financial years within the same 10-year period. The financial year denoted 01/02 corresponds to the period beginning 1 July 2001 and ending 30 June 2002. Black circles are the median of the posterior distribution, dark grey regions span the 25th to 75th percentile and light tray spans the 2.5th to 97.5th percentile. In years where there were no missing data, the red crosses show the loads computed using observed data only.

Figure 12. Posterior load estimates over 10 calendar years, where the following observation data were omitted from the original data: (a) all total suspended sediment (TSS) observations for 2001 (b) all discharge and TSS for 2001 (c) all TSS for 2010 and (d) all TSS and discharge for 2010. Black circles are the median of the posterior distribution, dark grey regions span the 25th to 75th percentile and light tray spans the 2.5th to 97.5th percentile. Red crosses show the loads computed from the original dataset (there are no red crosses in 2007–2009 because of actual missing observations in the original data). Observations. As shown in Figure 12, we were able to make annual load estimates within a plausible range for all of these missing data scenarios. Not surprisingly, estimates for these scenarios had substantially wider credible intervals than seen in Figure 11, but this was expected on account of the omitted data. Reassuringly, for all four scenarios in Figure 12, the load estimated from the complete observational data (shown with the red cross) was always within the credible interval for the load estimate and was generally close to the median of the posterior distribution for the load. Where measurements were not available, we found that the prediction was influenced heavily by the process model and the uncertainties placed on the understanding of that process. Consequently, we observe a change in the load estimates for
years where data were omitted (and potentially in other years as well) because parameters had been estimated from substantially different datasets. Not surprisingly, the credible intervals were wider when both TSS and discharge data were omitted than when TSS data alone were omitted, because the latter estimates relied heavily on two process models for prediction. It was also noteworthy that the omission of different parts of the dataset could subtly affect the width of the credible intervals for years where data were not omitted. For example, the credible interval for 2009 is slightly different for each of the four scenarios, because the estimates in these years are dependent upon the available data in other years to estimate process dynamics.

5. CONCLUSIONS

The highly variable nature of water quality data and flows makes quantification of the loads a challenging process. To be considered scientifically rigorous, load estimates should acknowledge the various sources of uncertainty in the underlying processes and the measurements used to support these processes. Doing so produces honest and defensible estimates for decision making. Estimating sediment loads without reporting errors and some measure of confidence is not only bad practice but can give false impressions to land managers and policymakers on the reliability of field measurements and how precisely a computer model can represent these environmental processes. We have demonstrated a statistically rigorous approach that can be used to assimilate hydrographic and water-quality data with process-based computer models using a BHM framework. To our knowledge, this is the first instance in which such methodology has been applied to estimating annual sediment loads. This is certainly the first application of such methods within a GBR catchment. It is our opinion that the BHM approach for estimating loads should become best practice for agencies in charge of reporting loads for the GBR. Adoption of such methods would draw strength from both models and data to statistically estimate loads based on all available scientific information in a rigorous manner.

Acknowledgements

We would like to acknowledge Scott Wilkinson for his contributions to the load monitoring methods used at Weany Creek, and to Aaron Hawdon and Anne Henderson for their management of the Weany Creek data collection, data processing and figure production. Data collection was funded by Meat and Livestock Australia, CSIRO and eWater CRC. We would also like to acknowledge the support and funding from Reef Rescue and CSIRO through the Water for a Healthy Country Flagship who made the development of this modelling framework possible. Funding support from CSIRO through the Julius Award is also gratefully acknowledged by the second author. Wikle acknowledges support from the US National Science Foundation grant DMS-1049093. We also acknowledge the interesting discussions around this topic with Noel Cressie and Mark Berliner.

REFERENCES


