CONSTRUCTED WETLANDS AS URBAN WATER QUALITY CONTROL PONDS - STUDIES ON RELIABILITY AND EFFECTIVENESS

Andrew Thomas, R. J. Morison, Philomena Gangaiya, A.G. Miskiewicz, R. L. Chambers and M. Powell

1Aurecon, 8/850 Collins St, Melbourne VIC 3008
Email: thomasa2@student.unimelb.edu.au
2Wollongong City Council, 41 Burelli St, Wollongong, NSW 2500
3University of Wollongong, Northfields Ave, Wollongong, NSW 2522
4Optimal Stormwater Pty Ltd, Level 5, 79 Victoria Ave, Chatswood, NSW 2067

Abstract: Constructed wetlands have come into widespread use as water quality control systems in urban areas. However published research into their water quality improvement performance has been sporadic and often contradictory. To address this, a cooperative study was initiated in 2009 by Wollongong City Council and the University of Wollongong to investigate the pollutant reduction performance of an existing water quality control pond. The pond was monitored for a period between 2009 and 2010 and a unique method for estimating constructed wetland performance was developed to address limitations found in other studies. This method incorporated automated sampling, high temporal resolution monitoring and standard least squares procedures to fit multivariate statistical models to estimate the pollutant reduction performance. The monitoring results were used to calibrate and validate a model which is able to quantitatively assess uncertainty. Results from this study suggest the method applied could be used as a standard method for estimating the pollutant reduction performance of other similar water quality improvement systems.

Key words: constructed wetlands, performance monitoring, urban development; water quality, pollutant mitigation, runoff, stormwater, water sensitive urban design WSUD, automated sampling, nitrogen, phosphorus.

INTRODUCTION

Nitrogen and phosphorus, along with sediments (suspended solids) have been identified as the top "Priority Pollutants" in NSW (DECC NSW 2009). Constructed wetland systems built to treat urban stormwater runoff are typically designed to reduce these pollutants in stormwater. Design reduction targets are load-based (average annual) and, depending on local conditions and environmental sensitivity, call for a 45-65% reduction in total nitrogen (TN), a 45-85% reduction in total phosphorus (TP) and an 80-90% reduction in total suspended solids (TSS) (DECC and CMA 2007, Landcom 2009a, MWC 2013, OEH 2013).

Nitrogen and phosphorus occur naturally in catchment runoff, and their presence in waterways is important for healthy biological activity. However, urbanisation typically results in an increase in these nutrients in receiving waters (Livingston 1990, Urbonas 2000, US EPA 2002, Gnecco, Berretta et al. 2005, Goonetilleke, Thomas et al. 2005, Egodawatta, Thomas et al. 2007, Farahmand, Fleming et al. 2007).

Increased nitrogen and phosphorus availability can have a significant impact on ecological processes in surface waters leading to reduced biodiversity, reduced resilience and, in some cases, complete system collapse (DECC NSW 2009). In particular, elevated levels can result in the eutrophication of a water body which can, under the right conditions, result in increased growth of aquatic plants including phytoplankton, cyanobacteria, macrophytes, seagrasses, and algae blooms (ANZECC and ARMCANZ 2000). This excessive growth can lead to a number of environmental and economic problems in surface waters, including aquatic fauna kills due to the release of toxins and deoxygenation of the water column, reduced recreational amenity, stock poisoning, reduced hydraulic conductivity / increased flood risk, altered and often reduced biodiversity and impacts on the provision of potable water supplies (ANZECC and ARMCANZ 2000, Osman Akan and Houghtalen 2003, DECC NSW 2009).

Changes in catchment hydrology due to urbanisation typically leads to an increase in peak discharges which increases suspended solid loads, particularly where construction works and / or unsealed roads have left soils exposed. These suspended solids impact on receiving waters in two ways. The first is physical and includes increased turbidity and smothering (Osman Akan and Houghtalen 2003, DECC NSW 2009). Increased turbidity reduces light penetration in the water column and this can have a number of impacts on aquatic organisms including sensory deprivation, reduced photosynthesis and reduced pathogen disinfection. Smothering of benthic habitat (e.g., seagrass beds) can also occur due to deposition of suspended solids when flow energies dissipate, especially where streams converge into larger water bodies (e.g., lakes, lagoons and oceans). Such deposition can also block pipes and channels, disrupting flow and potentially increasing flood risk (Duncan 2006). The second is the provision of a transport vector for other pollutants such as hydrocarbons, heavy metals, pathogens, organic matter and nutrients (particularly phosphorus) through their sorption to the particulate matter (Goonetilleke, Thomas et al. 2005, DECC NSW 2009). This relationship between suspended solids and other
pollutants has seen suspended solids used as an indicator for urban runoff pollution (Duncan 2006), and as surrogates for specific pollutants in specific catchments (Landcom 2009d).

Over the past five decades, much effort has gone into the science underpinning the use of “wet” (e.g., wetlands, ponds) water quality improvement systems that utilize natural processes to improve water quality (Kadlec and Knight 1996, Shutes 2001). By the 1990s, the use of these system for mitigating the environmental impacts of urban runoff on receiving waters had become popular (DLWC 1998). Since that time, the designs and efficiencies of these systems have evolved considerably. Despite these advances, the in-situ measurement of nutrient reduction performance still remains a challenge, resulting in uncertainty concerning the performance of these systems, particularly as they age (Goonetilleke, Thomas et al. 2005, Kadlec and Wallace 2008b, Ahialblame, Engel et al. 2012). There are two main reasons for this. The first is that there is currently no technology available that can directly measure nitrogen, phosphorus and suspended solids at the necessary sensitivity and temporal resolution (Jones 2008). The second concerns predicting exactly when a rain event will start and then taking enough samples and measurements to adequately capture the effects of the complex interactions between rainfall intensity and pollutant behaviour.

The challenge of estimating the in-situ performance of a constructed wetland was taken up in a collaborative University of Wollongong–Wollongong City Council project. The initial aim of this project was to estimate the capacity of water quality control pond (WQCP) “ROB1”, a simple form of constructed wetland, to reduce the loads of the priority pollutant” Total Nitrogen (TN), Total Phosphorus (TP) and Total Suspended Solids (TSS) (DECC NSW 2009). But, after a comprehensive review of the literature, it was not possible to identify a suitable and consistent method for this purpose (Thomas 2013).

The review indicated that the accuracy of many of these previous studies was questionable (some authors specifically admitted this); either because sampling frequency was insufficient or because the method of determining removal efficiency was likely to be misleading (Carleton, Gizzard et al. 2000, Kovacic, David et al. 2000, Farrell and Scheckenberger 2003, Tanner, Nguyen et al. 2005). For example, Dong et al. (2011) and Farrell and Scheckenberger (2003) report quite different load reductions for total nitrogen. This could be due to factors like wetland design, treatment water source, and climatic conditions. Unfortunately, variations in data collection (method and frequency), load calculations, and statistical procedures make such inter-study comparisons scientifically unsound. This has serious implications for published literature concerning the performance of constructed wetlands and the assumptions around the importance of these systems as water quality control measures. As a consequence of the above, the investigation into the performance of ROB1 (a 20 year old WQCP located south of Sydney) became as much about developing a quantitative method of estimating in-situ performance of constructed wetlands as it was about determining the pond’s pollutant reduction performance.

The aim of the study consisted of two parts. The first was to develop a quantitative method of estimating the capacity of a water quality improvement system like ROB1 to reduce total nitrogen (TN), total phosphorus (TP) and total suspended solids (TSS) in urban runoff. The second was to then apply this method to estimate the performance of ROB1 and use this information to infer its performance based on the conditions of the catchment and the pond at the time of the experimental phase of the work.

MATERIALS AND METHODS

Study site

The study site was located at Horsley Estate, a small satellite suburb nestled between Dapto and the escarpment of the Illawarra region of New South Wales (NSW) (Latitude 34°29’7.45”S, Longitude 150°46’19.64”E). The region is warm and temperate, with an annual average maximum daily temperature of 22.9°C and an annual average rainfall of approximately 1,100 mm. The study site consisted of a WQCP, “ROB1” and its catchment (Figure 1). The pond was originally constructed as a sedimentation basin during the construction phase of Horsley Estate in the early 1990s. This settling basin was subsequently converted into a WQCP in 1996 to protect downstream waters bodies from the impact of urban stormwater runoff. This conversion process involved the establishment of macrophyte vegetation, mostly Phragmites australis with other types of Poaceae and Typha orientalis (Figure 2).

The catchment area draining into ROB1 had a total area of 33.2 ha. The surface area of ROB1 was 1.14 ha, or just under 3.5% of the catchment area. The pond was designed to provide a permanent storage volume of 20.23 ML of water (Sinclair Knight 1994).

Data Collection

Data Acquisition

To meet the challenge of capturing appropriate data at sufficient resolution to estimate water quality improvement performance, five water quality monitoring stations (WQMS) were established at ROB1, i.e., four at each of its inlets (H1, H2, H3 and H4) and one at the outlet (H5) (Figure 2). Each station consisted of an auto sampler, flow monitor (MACE FloPro Series 3) turbidity sensor (Campbell Scientific OBS-3+), conductivity and temperature sensor (Campbell Scientific CS547A Conductivity and Temperature Sensor). A rain gauge (Tipping Bucket Rain Gauge TB4) was also installed at WQMS H1 (Figure 3).
Establishing these WQMS made it possible to monitor water flow (L/s), turbidity, conductivity and temperature at high temporal resolution whilst allowing for sample collection to occur at any time of the day or night based on predetermined programming. Each WQMS was time synchronized to allow direct comparison of data at each of the inlets and the outlet. Collectively, these five WQMS constituted the “ROB1 Performance Monitoring System” (ROB1 PMS).

**Monitoring and Sampling Regime**

Water quality monitoring and sampling needed to provide the temporal resolution necessary to account for the high variability in rainfall intensity and uncertainty typically associated with individual rainfall events in order to estimate load balances reliably within the constraints of technological and budgetary realities. Since the CR800 data loggers were capable of
Figure 2: Schematic of ROB1 (Hopkins and Yassini 2006). Note: GPT – Gross Pollutant Trap, CDS – Continuous Deflective Separation stormwater pollution trap.

Figure 3: One of five similar water quality monitoring stations set up to monitor the inlets to, and outlet of, ROB1
storing large amounts of data, it was possible to take readings for flow, turbidity, conductivity and water temperature at five minute intervals for the full length of all rain events monitored. TN, TP and TSS were determined by taking water samples and sending them for laboratory analysis. These water samples were collected using auto-samplers that were limited to 24 samples. To compensate for this limited number of water samples, each inlet auto-sampler (i.e., H1 to H4) was programed to take samples every 10 minutes during the first hour (starting at time zero) and every half hour thereafter during a rain event. The WQMS at the outlet, H5, was programmed similarly, except that after the 7th sample, remaining samples were taken every hour thereafter. This difference was due to the outlet being far more predictable in terms of flow rate and pollutant concentration due to the homogenising influence of the pond and the controlled release of water due to the design of the outlet structure.

Six separate sampling events were captured by the ROB1 PMS between October 2009 and September 2010. Most of the rain periods sampled were discrete events that began and ended with clear start and end points. The only exception was Event 2, which occurred during an extended wet period during February 2010.

It is noted that during the experimental period (2009-2010), annualised and monthly rainfall totals from nearby Bureau of Meteorology weather stations 68000 and 68022 indicates that rainfall for the experimental period was only slightly above average. Based on monthly averages, the individual events were not particularly unusual for the study site.

Estimating Pollutant Load Balances

Overview

A three-stage process was developed to estimate pollutant load balances. The first involved collating data from all five monitoring stations and the subsequent determination of derivative parameters including “station number”, “rain event” and “cumulative flow”. The second stage involved the construction of predictive models to predict TN, TP and TSS concentrations using the high resolution monitoring data. The third stage involved using the predicted TN, TP and TSS data along with the flow data to estimate pollutant loads. The first and third stages were undertaken using Microsoft Excel (Version 14.0.6112.5000) and the second stage was undertaken using JMP Statistical Software (version 9.0.2).

While a total of six events were monitored during the experimental phase, only four events were ultimately used to calculate pollutant load balances; Event 1, Event 2c, Event 5 and Event 6. Events 3 and 4 were excluded due to incomplete data sets resulting from technical and operational problems that occurred with the field equipment during these Events (Table 1). Similarly, Event 2 as a whole could not be used; however, due to its size, it was both possible and beneficial to delineate and extract a subset of data from Event 2 to calculate pollutant load balances, i.e., “Event 2c”.

![Figure 4: Illustration of the assignment of sub-events to an event using Event 1 as an example. Red dots show when water samples were taken and the blue line, “Flow”, is based on flow readings taken every 5 minutes at WQMS H1 during Event 1. Note the difference in the nature of rainfall and subsequent flow patterns for each sub-event.](image-url)
Data Collation

Data was collated into individual sampling “events”. Delineation of a given event was based on rainfall, where a period greater than 48 hours without any precipitation was considered to be the minimum time lapse between two events. Sub-events were delineated based on an examination of hydrographs of an event and were assigned based on short periods of no measurable flow indicating short gaps in rainfall (Figure 4). Sub-events were derived to account for the potential effects associated with rainfall intensity changes and other unforeseen consequences of disrupted rainfall/surface runoff flows. “Cumulative Flow” was added to the data set to take into account potential “wash-off” effects, i.e., the reduction and change in pollutants that may be expected to occur as a rain event continues (Duncan 1995, Duncan 1999, Duncan 2006). This was derived by simply multiplying up each five minute flow reading (L/s) by 300 seconds and adding each to the previous values.

Establishing Predictive Models

Two models were established to predict TN, TP and TSS, i.e., one for the inlets, referred to as the Inlet Predictive Model (IPM) and one for the outlet, the Outlet Predictive Model (OPM). Two models were needed to account for the substantially different environments involved, drainage catchments (IPM) and pond (OPM).

A “Modelling Data Spreadsheet” was created in Excel by extracting Laboratory results along with their corresponding monitoring results, calculated cumulative flow and binary parameters. This Modelling Data Spreadsheet was then imported into JMP. Before fitting the model, variables with highly skewed distributions were transformed to their natural log (i.e., turbidity, conductivity, flow rate, cumulative flow, TN concentration, TP concentration and TSS concentration). This reduced the instability associated with modelling highly skewed data sets by ensuring that more normally distributed datasets best suited to model-fitting via standard least squares were used in the analysis.

To establish the IPM, a multivariate linear regression model was then fitted via standard least squares to Log TN, Log TSS, and Log TP (response variables) for imported data from WQMS 1 through WQMS 4 inclusive using the following explanatory variables: log turbidity, log conductivity, log flow rate, log cumulative flow, water temperature, rain event, and Sub event. Station (i.e., each WQMS was included to account for idiosyncratic affects associated with each of the inlets and catchments).

Modelling included both main (direct) linear effects, as well as second order interactions (indirect effects) to account for potential non-linearity in the relationships between explanatory and response variables. Screening was applied to the model in order to identify and remove non-significant terms. Model regression diagnostics were used to check model performance and remove outliers.

The OPM was constructed in effectively the same way except that data from only one WQMS was used (i.e., H5) and sub-events were excluded as the engineered outlet structure controlled outlet flow to the extent that sub-events were not discernible.

RESULTS

Data Collection

A key aim of this study was to develop a quantitative method of estimating the capacity of a water quality improvement system like ROB1 to reduce TN, TP, and TSS in urban runoff. A crucial part of this aim was to collect data at sufficient temporal resolution to adequately resolve the complex behaviour of pollutants within the catchment and pond systems. Table 1 provides a summary of the outcomes of the implementation of the ROB1 PMS installed to achieve this objective. The green, yellow and red dots are a “stop light” system used to convey a visual qualitative assessment of how each station performed. Green means operated as per design (good). Yellow means operational issues resulted in moderate reduction in data capture quality and, whilst data was suitable for statistical modelling, it was not suitable for load estimations. Red equals serious operational issues have led to data loss to an extent such that it is not suitable for either load estimations or statistical modelling.

Model Diagnostics

Model diagnostics played two important roles in this study. The first was the provision of the capacity to quantify the confidence in the IPM and OPM and, therefore, the pollutant load reduction performance of ROB1. The seconds was as a means of “proof” that the methodology applied in this study is likely to be a true reflection of the actual performance of ROB1 in the absence of validation studies and any means of direct validation (no technology available to do this).

Model diagnostics are summarised in Table 2. The R2 values for Log TN and Log TP indicated that both the IPM and the OPM were able to explain a high proportion of the variation for these variables. The analysis of variance outputs for each model reveal high F ratios and degrees of freedom that are well below the number of observations used to construct the models, indicating that the chance of the relationships found being merely coincidence is extremely unlikely. Predicted plots showed strong linearity and residual plots showed no discernible patterns indicating that no important correlations between response variables and explanatory variables had been missed.
Table 1: Summary of the performance of the ROB1 PMS

<table>
<thead>
<tr>
<th>Event</th>
<th>Date/Locations</th>
<th>Samples collected</th>
<th>Details</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event 1:</td>
<td>25-26 Oct 2009 (42.5 mm in 21.5 hours)</td>
<td>WQMS H1 = 24, WQMS H2 = 24, WQMS H3 = 24, WQMS H4 = 24, WQMS H5 = 24, Overall = 120</td>
<td>All WQMS worked as intended, however the latter half of the event was not well captured in terms of water quality samples due to logistical limitations of the methodology. This was corrected by changing the sampling program to increase time between samples and by developing a bottle collection and replacement regime to facilitate smooth bottle exchange during an event.</td>
</tr>
<tr>
<td>Event 2:</td>
<td>Early to mid-Feb 2010 (118 mm over six days)</td>
<td>WQMS H1 = 48, WQMS H2 = 96, WQMS H3 = 63, WQMS H4 = 78, WQMS H5 = 152, Overall = 437</td>
<td>The first 24 hours of this extended event was missed due to a program coding error that resulted in the entire ROB1 PMS failing to initiate. This was further complicated by the auto sampler at WQMS H1 stopping after the collection of only six samples. The first 24 hours of lost data was rectified by reloading the program with the coding error corrected. The cause of the problem with WQMS H1 was unclear at the time, but was temporally rectified by resetting the auto sampler and re-starting. Unfortunately, this issue re-occurred as detailed for Event 3 below.</td>
</tr>
<tr>
<td>Event 3:</td>
<td>28 Feb – 2 Mar 2010 (28.5 mm in 34.25 hours)</td>
<td>WQMS H1 = 7, WQMS H2 = 0, WQMS H3 = 8, WQMS H4 = 15, WQMS H5 = 24, Overall = 54</td>
<td>Two unrelated problems lead to a marginal event in terms of the usefulness of data for load reduction calculations, however data was still viable for statistical modelling. H1 again suffered a distributor arm failure after the collection of six samples, preventing further collection of samples for the remainder of the event. Also, H2 suffered from a FloPro software freeze resulting in no flow data and, consequently, no water samples collected at WQMS H2 for the entire event. The re-occurrence of the problem with WQMS H1 was traced back to degraded seals associated with the auto sampler controller casing which allowed moisture to come into contact with the auto-sampler PLC motherboard. The cause of the freezing of the flow monitoring software at WQMS H2 was not clear. It was thought that some sort of power surge (e.g. lighting strike) may have been the cause. Due to this uncertainty, no specific solutions were implemented other than rebooting the FloPro system at WQMS H2.</td>
</tr>
<tr>
<td>Event 4:</td>
<td>30 Mar – 5 Apr 2010 (73 mm in 6 days and 23.75 hours)</td>
<td>WQMS H1 = 35, WQMS H2 = 69, WQMS H3 = 14, WQMS H4 = 0, WQMS H5 = 117, Overall = 235</td>
<td>Event 4 was not suitable for load estimations and could only be used for modelling for two reasons. The first was due to WQMS H4 having no auto-sampler as its auto-sampler had been relocated to H1. The second was due to another FloPro software freeze, this time at WQMS H3. Because this problem was caught early in the event, data from H3 was still useful for modelling but not for load calculations. Again, the root cause of the freezing of the FloPro software could not be determined (possibly heat), however the problem was managed successfully for future events by implementing pre-start checks that included ensuring the FloPro software was working as intended at each WQMS.</td>
</tr>
<tr>
<td>Event 5:</td>
<td>2 – 4 September 2010 (25 mm in 2 days and 13.5 hours)</td>
<td>WQMS H1 = 13, WQMS H2 = 27, WQMS H3 = 5, WQMS H4 = 28, WQMS H5 = 74, Overall = 147</td>
<td>All stations fired as per design, however H3 failed to collect the second round of samples due to a peristaltic pump tube failure (split) caused by wear due to the constant action of the peristaltic pump rollers on the tubing. Otherwise, H3 performed as per design. This problem was solved simply by replacing this tubing.</td>
</tr>
<tr>
<td>Event 6:</td>
<td>14 – 15 Sept 2010 (25 mm in 1 day and 17.25 hours)</td>
<td>WQMS H1 = 24, WQMS H2 = 19, WQMS H3 = 24, WQMS H4 = 24, WQMS H5 = 35, Overall = 126</td>
<td>A near perfect event, representing a culmination of the lessons learnt about the ROB1 PMS during the data collection phase. The only issue that impacted on this event was multiple zero flow readings for WQMS H3. The cause of these zero readings was put down to the intensity of the rainfall during Event 6 which produced periods of very high flows, and the proximity of the flow meter to the trash rack at inlet H3 which created high turbulence during these high flow periods.</td>
</tr>
</tbody>
</table>
Figure 5: Histogram of rain events size distribution for all recorded 24 hour rainfall records between 1930 and 2011, including cumulative percentage frequencies (Source: BoM Albion Park Post Office weather station #068000).

Rainfall Event Size Frequency Histogram

Rainfall / 24 hours - 1930-2011. Albion Park, BoM Weather Station # 068000
Diagnostics for TSS for both models were not as strong, particularly for the OPM. Whilst F ratios for both models were reasonable, and the R2 value for the IPM was relatively high, the R2 value for the OPM was poor (Table 2). Due to the poor R2 result for TSS, pollutant load reduction estimates could not be reliably calculated as the uncertainty associated with such results would simply be too high to be useful.

**Pollutant Loads Reductions**

The results for each event are summarized in Table 3. The estimated reduction of TN and TP by ROB1 during Event 1 was 69% for TN and 43% for TP. Stormwater runoff entering the pond during Event 1 resulted in the displacement of approximately half of the volume of water present in ROB1 immediately prior to its commencement. Retention of TN and TP were also recorded for Event 5 (71% and 75% respectively), with a slightly lower pond volume turnover of 40%.

In contrast, the export of pollutants was observed for Event 2c and Event 6. Event 2c resulted in the export of both TN (-39%) and TP (-9%), whereas Event 6 recorded an export of TN only (-58%). This coincided with the amount of stormwater runoff entering the pond exceeding the estimated pond volume at the start of the event for both of these events (Table 3).

In order to illustrate the performance of ROB1 using the results presented in Table 3, a histogram (Figure 5) has been constructed using historic rainfall data from the local area. Superimposed onto this histogram is each sampling event for which load reduction estimates were calculated according to each event’s recorded rainfall. This histogram provides a backdrop of rainfall frequency for the local area enabling a comparison of each sampling event relative to the historic rainfall patterns associate with the study site.

**DISCUSSION**

**General**

Results reveal that on an event basis ROB1 is likely to be retaining pollutants and it may do so for approximately 80% of rain events likely to be experienced at its location, depending on antecedent conditions. For larger events, particularly those that result in complete pond turnover, performance appears to wane dramatically, leading to the export of TN and/or TP. Results also reveal that while validation of the experiment is not possible due to the current technological limitations of water sampling, statistical diagnostics strongly suggest that the correlations identified were not random; hence the calculated load balances are likely to reflect actual performance.

**Table 3: Priority Pollutant load estimation calculation results for Event 1, Event 2c, Event 5 and Event 6.**

<table>
<thead>
<tr>
<th>Event</th>
<th>Inflow Load (g)</th>
<th>Outflow Load (g)</th>
<th>% Reduction</th>
</tr>
</thead>
<tbody>
<tr>
<td>Event 1</td>
<td>13,485</td>
<td>4,143</td>
<td>69%</td>
</tr>
<tr>
<td>Event 2c</td>
<td>24,005</td>
<td>2,881</td>
<td>-39%</td>
</tr>
<tr>
<td>Event 5</td>
<td>33,283</td>
<td>8,416</td>
<td>-71%</td>
</tr>
<tr>
<td>Event 6</td>
<td>47,393</td>
<td>11,938</td>
<td>-58%</td>
</tr>
</tbody>
</table>

*Estimated pond volume at the start of each event was estimated using bathymetric data as described in Thomas (2013). Estimated pond turnover was determined by dividing the estimated pond volume at the start of each event by the total outflow.*

**Table 2: Statistical Diagnostics for the IPM and OPM for each priority pollutant: TN, TP and TSS.**

<table>
<thead>
<tr>
<th>Prediction Model</th>
<th>Observations</th>
<th>R²</th>
<th>F ratio</th>
<th>P value</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Inlet Predictive Model (IPM)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degrees of Freedom = 64</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TN</td>
<td>660</td>
<td>0.89</td>
<td>73.4</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>TP</td>
<td>688</td>
<td>0.81</td>
<td>39.7</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>TSS</td>
<td>662</td>
<td>0.71</td>
<td>23.2</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td><strong>Outlet Predictive Model (OPM)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Degrees of Freedom = 20</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>TN</td>
<td>425</td>
<td>0.83</td>
<td>96.3</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>TP</td>
<td>418</td>
<td>0.90</td>
<td>175.6</td>
<td>&lt;0.0001</td>
</tr>
<tr>
<td>TSS</td>
<td>418</td>
<td>0.54</td>
<td>23.3</td>
<td>&lt;0.0001</td>
</tr>
</tbody>
</table>
Data Collection

The ROB1 PMS was installed to overcome shortcomings in the methods implemented in other studies. Due to the reliance of the ROB1 PMS on integrated mechanical and computer based technologies, this decision brought with it higher costs and complexity relative to other sampling techniques identified in the literature. This complexity lead to some partial system failures which did result in a reduction in the quality of the data collected for some of the sampling events. However, most of the system failures were avoidable and were often down to simple human error (e.g. coding error; Event 2) or operator inexperience (e.g. peristaltic pump tube failure; Event 5). When the ROB1 PMS operated as intended, data capture was of a high quality rarely seen in the published literature and subsequent modelling diagnostics demonstrate that data capture was of sufficient quality to produce robust estimates for both TN and TP.

Confidence in the Predictive Models (IPM and OPM)

Both the IPM and the OPM demonstrate strong capacity to explain the variation associated with TN and TP based on the explanatory variables used and the model rules selected. For each model, response coefficients (R2) are quite high (0.81 – 0.90) and the associated analysis of variance outputs indicate that the chance that the relationships between explanatory variables and response variables being random (i.e., mere chance) are remote. Given the environment in which this study was undertaken and the complicated interactions likely to be occurring, the R2 values achieved and the strong analysis of variance results are surprisingly robust. Consequently, both models were considered acceptable for estimating Log TN and Log TP (and the subsequent derivation of TN and TP) for this study.

The strong results for TN and TP were not repeated for TSS. While analysis of variance outputs for both models for Log TSS were quite reasonable, the response coefficient (R2) was particularly poor for the OPM (0.54), meaning predictions for TSS were too uncertain to be of any real use for performance assessment. Root cause analysis determined that the most likely cause was not the design of the models, but rather TSS laboratory detection limits as evidenced by the high number of non-detections reported, particularly at the outlet. Hence, it is likely that if a more sensitive TSS analysis method was used diagnostics would have been much stronger, possibly mirroring the strong results achieved for TN and TP.

Pollutant Reduction Performance of ROB1

ROB1 appears to have the capacity to achieve load reductions for both TN and TP for around 80% of all rain events likely to occur during its lifecycle. Results also suggest that ROB1 may be achieving pollutant reductions in the order of 70% or better for both TN and TP for rain events of 25.5 mm or less, accounting for approximately 60% of all potential rain events likely to be experienced by ROB1 during its lifecycle. These reductions are consistent with reductions reported in other published studies (Carleton, Gizzard et al. 2000, Farrell and Scheckenberger 2003, Fisher and Acreman 2004, Terzakis, Fountoulakis et al. 2008, Ko, Chang et al. 2010).

However, as also reported in the other published studies (Kovacic, David et al. 2000, Fisher and Acreman 2004, Tanner, Nguyen et al. 2005), the export of TN and/or TP was also observed for Event 2c and Event 6, both of which also saw pond turnover exceed 100%. This suggests that the capacity of ROB1 to retain TN and TP begins to wane significantly during rain events that create enough surface runoff to completely displace the original pond volume at the start of an event. Performance also seems to have been affected by antecedent rainfall, where longer periods of dry weather between events appear to favour the pollutant retention.

Other issues that may have been reducing ROB1’s capacity to retain pollutants, but cannot be ascertained directly using the data collected in this study, relate to the poor design of ROB1 relative to current best practice (Water by Design 2006). This includes the lack of high flow bypass, the relative position of some pond inlets to the outlet (increased risk of short circuiting) and a lack of effective “pre-treatment”, i.e. pre-pond sediment treatment and gross pollutant trapping.

The impact of poor design may have been, in particular, a factor for reduced retention of TP and export of TN during Event 6. This event was characterised by intense, high volume rainfall which appears to have effectively swamped ROB1’s capacity to treat stormwater despite a reasonable gap in antecedent rainfall leading up to that event. In contrast, the export of both TN and TP for Event 2c appears to be function of reduced retention time due to the consistent and recent antecedent rainfall leading up to Event 2c.

CONCLUSIONS

The results suggest that ROB1 is achieving reductions in TN and TP on an event basis, and that reduction for both TN and TP could be occurring for around 80% of the events the pond is likely to be exposed to during its lifecycle. However, for the remaining 20% of the more extreme events, potential exists for the export of pollutants. This, combined with low flow loads between events, suggests that the reported TN and TP load reductions achieved by ROB1 on an event basis may be over-estimating the overall performance of ROB1 in terms of its capacity to reduce TN and TP loading.

The results also suggest that ROB1’s design could be impacting negatively on its performance capacity. These design flaws include inadequate gross pollutant traps, no sediment traps at the pond inlets, no high flow bypass, and the relative close proximity of some of the pond inlets to the outlet, particularly H1 and H4. These design flaws reduce ROB1’s capacity to treat urban runoff by increasing the potential for the re-suspension of fine sediments and organic detritus, and increase the
risk of short circuit and swamping of the overall water quality improvement system during large, more extreme rain events. It is noted that such flaws would be excluded from modern systems designed according to best practice (URS 2004, Waterways 2006).

If validated, the method applied in this study could be used to quantitatively assess the performance of other similar stormwater quality control measures. Benefits of this methodology would include a robust means of comparing results between studies, assessing the performance of similar infrastructure as they age, assessing different treatment measure designs, and determining the actual performance of infrastructure following instillation which would benefit consent authority assessment of legislated systems and help to calibrate modelling software (e.g. MUSIC).

Given the strong statistical diagnostics presented in the results, it is recommended that further studies be carried out using the method described herein to validate findings, both at ROB1 and at other treatment systems at other locations with varying climates and catchment features. It is also recommended that the modelling undertaken for this study be more rigorously tested using improved water quality sampling technology as it becomes available. While such technology remains elusive, it is noted that recent innovations in high temporal resolution nutrient monitoring may see this situation improve (Wild-Allen and Rayner 2014). Full details of all the work undertaken and results obtained for this study are detailed in Thomas (2013).
REFERENCES


