Characterising Urban Pollutant Loads

Matt Francey B.Ed, B.Eng(Hons)

Thesis submitted in fulfilment of the requirements of the degree of Doctor of Philosophy

Department of Civil Engineering

Monash University

February 2010

Abstract

The effect of urban runoff on aquatic ecosystems, the recreational and other functions of waterways is of concern in most, if not all major urban centres. Addressing the impacts of poor stormwater quality is a regular function of water and environment managers worldwide. A component of this function is the accurate and efficient modelling of the pollution sources, in order to both enable the best understanding of the processes involved and to facilitate the most cost effective treatment options.

Stormwater quality modelling has had a chequered history. Despite consistent attention over the past 30 years there is still little agreement of modelling approaches, or even the primary processes involved in pollutant generation.

The current study builds on past work which has suggested that rainfall intensity is a major driver, and has collected comprehensive monitoring data of a quality that reduces uncertainty in the results to a level where confusing factors are minimised. It also examines a number of catchments from a consistent meteorological region.

The study resulted in short interval rainfall, runoff and pollutant data being collected for over 250 events from seven catchments. TSS, TP and TN were most commonly collected but metals, nitrogen speciation and other pollutants were also collected for many sites. Based on a review of five data sets collected previously, both in Australia and internationally, the monitoring program aimed to collect enough events to minimise uncertainty related to the recorded variation in EMC values. Pollutant data collected showed significant differences to that previously collected in the region, with TSS occurring at concentrations approximately half of those recorded in past studies. Uncertainties due to data collection in pollutant and flow data confirm previously reported figures of approximately 30% when applied to prediction of loads.

To compare the principal factors behind pollutant generation, the study tested a simple two-parameter model for calculating event loads of total suspended solids (TSS), total phosphorus (TP) and total nitrogen (TN). Rainfall intensity, total rainfall, total flow volume and runoff intensity are tested as explanatory factors and a clear preference is shown for rainfall intensity, particularly for TSS and TP which have a greater proportion of their total mass in particulate form. It was found that the two-parameter model based on rainfall intensity typically explains about 90% of the variation in event loads at a site. A verification exercise was undertaken to assess the model's ability to predict catchments not used for calibration and the model predicts both total loads (average difference between observed and predicted loads of less than 20%) and loads for each event (average co-efficient of efficiency for TSS of approximately 0.7). It also examined the sensitivity of the model to its two parameters, and to the rainfall time step. Despite its acknowledged correlation with flow, rainfall intensity over short timesteps is considered to be the main driver of pollutant mobilisation, and provides a practical means of predicting pollutant behaviour using readily available data.

The model also satisfactorily predicts the within-event behaviour of pollutants, both loads and concentration, when the flow lag time is taken into account. Prediction of the lag time is identified as the major challenge in short timestep prediction of pollutants.

The uncertainty of the model inputs is considered and quantified, both for systematic and random errors. Error in rainfall data is the main factor in the total model error, and creates uncertainties of at least 30%.

The study confirms previous research, which has suggested rainfall intensity as the main driver behind pollutant generation, at least for TSS and associated pollutants. It also provides a simple to use, easy to calibrate model that could be used immediately by practitioners for pollutant prediction.

Notation

µg	microgram
BOD	Biochemical oxygen demand
Cd	Cadmium
COD	Chemical oxygen demand
CRC	Co-operative research centre
CRCCH	Co-operative research centre for catchment hydrology
CSIRO	Commonwealth Scientific and Industrial Research Organisation
Cu	Copper
E	Coefficient of efficiency
EMC	Event mean concentration
EPA	(Victorian) Environment protection authority
kg	kilogram
l	litre
m	metre
MUSIC	Model for Urban Stormwater Improvement Conceptualisation
NH_4-N	Ammonia nitrogen
NO_3-N	Nitrate nitrogen
NPS	Non point source
NURP	Nationwide Urban Runoff Program
Pb	Lead
Q	Flow
QTN	Flow rate of total nitrogen (g/s)
QTP	Flow rate of total phosphorus (g/s)
QSS	Flow rate of suspended solids (g/s)
R^2	Coefficient of determination
s	second

SMC	Site mean concentration
SWMM	Storm water management model
TN	Total nitrogen
TOC	Total organic carbon
TSS	Total suspended solids
TP	Total phosphorus
W	Washoff coefficient
WSUD	Water sensitive urban design
USA	United States of America
USEPA	United States environmental protection agency
Zn	Zinc
µEMC	Mean of the event mean concentrations
σEMC	Standard deviation of the event mean concentrations

Acknowledgements

I would firstly like to thank my supervisors; Ana Deletic, Hugh Duncan and Tim Fletcher. It has been particularly difficult trying to balance the competing demands of research, a full time job and an (expanding) family during the past years and they have shown patience and understanding throughout.

Peter Polesma and Justin Lewis did much of the hard work in helping to set up the monitoring program and then manage the collection of the data. David McCarthy provided valuable insights and constructive criticism at various points during the study.

EPA Victoria provided funding for the monitoring program through their Victorian Stormwater Action Program. Sean Moran was particularly important in instigation the work.

Finally Melbourne Water and my family supported me to undertake this study as a part of my employment and life respectively.

Declaration

This thesis contains no material accepted for the award of any other degree or diploma in any university and, to the best of the authors knowledge and belief, contains no material previously written or published by another person except where due reference is made in the text. The length of the thesis is less than 10,000 words, exclusive of figures, tables, references and appendices.

Signed:

1 INTRODUCTION

Urbanisation and associated activities have generally resulted in a decrease in the health of natural waterways. At least three processes are involved: changes in both the amount and timing of flows, the release of concentrated pollutants from specific sources and the generation and efficient transport to receiving waters of diffuse pollution. The deliberate or accidental release of pollutants, called point source generation, has generally been decreasing in modern cities, due to increasing environmental and health standards. The effect of flow changes and diffuse pollutants are still prevalent, however, and are increasingly coming under scrutiny.

In urban areas, impervious surfaces such as roofs, roads and pavements severely restrict, or completely prevent rainfall from reaching underlying soil. They also provide surfaces where various pollutants generated from anthropological activities such as metals, nutrients and hydrocarbons accumulate. Rainfall strikes these surfaces and is transferred in a hydraulically efficient manner through channels and pipes before discharge into receiving waters, such as lakes, streams or marine waters.

Descriptions of these effects of urbanisation are common in engineering and scientific literature (Duda et al. 1982; Sharpin 1993; Wong et al. 2000; Walsh et al. 2001; CRC for Freshwater Ecology 2004; Hatt et al. 2004; Walsh et al. 2004). The main effects include:

- Increased peak flows during storm events
- Increased volume of surface runoff
- Less attenuation of storm flows resulting in shorter periods of stormflow
- Reduced infiltration to subsurface flow and groundwater
- Increased loads of pollutants, increased concentrations of some pollutants such as heavy metals
- Changes in groundwater levels (can be either higher or lower due to the competing effects of changes to evapotranspiration, reduced infiltration, leakage from distribution systems and excessive water use)
- A decrease in the water quality of receiving water bodies
- A decrease in the ecological health (based on species richness) of waterways

These impacts affect both riverine and coastal systems (CSIRO 1996). In cities like Melbourne, stormwater has been identified by scientists and water management agencies as one of the major threats to coastal values. This has resulted in major expenditure on stormwater treatment. For instance the regional drainage authority in Melbourne owns and operates over 130 stormwater treatment wetlands (www.melbournewater.com.au), and operates a stormwater offsets scheme where developers must either treat their stormwater on-site, or pay to treat it elsewhere in the catchment.

1.1 Modelling of stormwater pollution, why is it important?

The modelling of urban stormwater can be split into two components, the simulation of the hydrology and the simulation of the pollutants. The hydrological component has a long history, based on flood modelling, and is therefore relatively well accepted in engineering research and practice. Although by no means an "exact science", the modelling of rainfall – runoff relationships in urban areas is widely understood (and used) by practitioners (Engineers Australia 1997). There are a wide range of models, of various complexities, that can consistently make predictions with sufficient accuracy for a given use.

The quality aspect of modelling, however, is still a "poor cousin", supported with far less local data and is generally much less well accepted. One major reason for this is that flooding produces immediate, obvious consequences that have direct economic and potentially life threatening impacts. Another is that hydrological processes are relatively consistent compared to the generation of pollutants. The effects of poor water quality however are both longer-term and the implications less direct in economic or health terms. The stochastic nature of urban pollutant concentrations (Duncan 1999) and the difficultly in accurately monitoring the loads delivered (Ahyerre et al. 1998) have also hindered the modelling of urban runoff.

Grayson (1992) gives two principal reasons for modelling:

1. Increased understanding of physical systems
2. As a predictive tool

Both of these reasons are particularly relevant to stormwater quality modelling. In contrast to modelling of flows, even the primary variables responsible for pollutant generation from urban catchments are poorly understood (Vaze and Chiew 2003). Predictions of stormwater quality are of most value in the planning and design of preventative measures. Management measures, such as the setting and assessment of targets for environmental performance, are dependent on an accepted, efficient and widely understood method of measurement. Modelling is the most cost and time effective method of achieving this goal. More efficient modelling, in terms of accuracy, therefore has a direct relationship to the design and environmental performance of stormwater management systems.

Structural stormwater treatment measures, which are becoming increasingly popular as a management tool, require the accurate assessment of where and when the main pollutant loads are carried – are pollutants concentrated in a particular part of the event or is there a uniform distribution? Are the majority of the loads in a system carried by large or small, intense or long, sustained events? How do designs need to differ depending on the pollutant targeted for example?

Accurate continuous modelling of both runoff and pollutant generation may provide efficiencies in terms of size, location and design of treatment measures and reduce practitioners' dependence on commonly used "rules of thumb". Over time, water sensitive urban design (WSUD) measures are also moving from being solely large, "end of pipe" systems to more distributed systems with elements designed to treat runoff as close as possible to its source. Sizing of elements is particularly important in this case, as space and maintenance constraints are often severe.

Development of robust yet simple-to-apply stormwater quality models could thus play a significant role in advancing the management of urban stormwater, and reducing its impacts on receiving waters. This thesis aims to make a contribution to this aim by developing a simple method, based on the primary factors behind pollutant generation that could be used to predict stormwater pollutant loads and improve the design of stormwater treatment devices.

1.2 Chapter outline

Chapter Two provides an overview of the previous literature on the topic of pollutant description and modelling from urban catchments. It describes work on the major physical processes involved and attempts to model those behaviours. The reasons for model development are discussed. It also provides an introduction into published work regarding the uncertainty of stormwater modelling. Reviews of relevant literature are also included in specific chapters. Chapter Two concludes with a synthesis of the central research questions addressed by this study.

Chapter Three presents descriptions of the catchments and data sets used to scope this study, and the development of a monitoring program specifically to test the study hypotheses. A detailed methodology for minimising monitoring uncertainty is also presented.

Chapter Four presents the initial results from the monitoring program. This includes the use of statistical methods to compare observed pollutants and flow with catchment and hydrological characteristics. It also compares observed pollutant behaviour with various hydrological descriptors, in order to provide insights into processes affecting pollutant concentrations.

Chapter Five proposes a model for event load prediction based on the literature review and data collection. A calibration/verification exercise is undertaken using the proposed model. The chapter examines the uses of the model for different pollutants both within and between different catchments. It also examines the sensitivity of model parameters to different model conditions.

Chapter Six examines the model performance at short timesteps within individual storm events. The concepts of lag and attenuation are examined. The chapter also provides an examination and discussion of the degree of model complexity which may be justified and appropriate for practical application.

Chapter Seven calculates the various uncertainties associated with the model. The model uncertainty is compared to the uncertainty associated with measurement of pollutant concentrations.

Chapter Eight presents the conclusions together with strengths and weaknesses of the evidence and discusses further research questions that have emerged during the study.

2 LITERATURE REVIEW AND DEFINITION OF THE RESEARCH PROBLEM

2.1 Stormwater as a non-point source of pollution

Flow in stormwater systems is largely derived from rain through surface runoff, but may receive contributions such as groundwater infiltration, over-application of water from anthropological sources (e.g. excessive watering of gardens) or leaks from sewage or water supply pipes. The surface runoff component is usually considered to be the greatest contributor to pollution loads, although in some situations and for some parameters dry level flows may be significant (McPherson 2002). An extreme example of this is found in many North American and European cities where sewage and stormwater are transferred together in “combined systems”, pollutants such as nutrients and pathogens may be very high during low flows (Deutsh and Hemain 1984). Australian cities are almost exclusively serviced by separate stormwater and sewage systems; therefore separate systems will form the basis of this thesis.

These systems interface with natural waterways, usually in the form of direct discharges into rivers, streams, wetlands or bays. The impact of stormwater on receiving water bodies has been extensively described (Jones 1986; House et al. 1993; Walsh, Fletcher et al. 2004). The health of receiving water bodies consistently declines with urbanisation in their catchments, even at very small levels. Large water bodies such as lowland rivers, lakes or bays are susceptible to long-term loads of pollutants, while smaller streams are impacted mostly by a combination of pollutant concentrations and flows (Hatt et al. 2004; Clark 2006);

Understanding of stormwater pollutant characteristics has greatly increased since the 1960s. Weibel et al. (1964) presented data from a group of early studies into urban runoff, which details BOD concentrations similar to secondary treated sewage, solids comparable to raw sewage and high nutrient levels. Pollutant loadings from urban stormwater have consistently been found to be similar to sewage with respect to solids, metals and hydrocarbons (Ellis 1986). Major French and US studies in the early 1980s reinforced the understanding of stormwater runoff as a significant polluter of waterways in urban areas (Deutsch and Desbordes 1981; Castaldi 1983).

A United States study, which was used to calibrate the US Environment Protection Agency’s NPS (Non Point-Source) model, described the generation of COD (chemical oxygen demand), lead, zinc, total phosphorus and total nitrogen (Hartigan et al. 1978). It showed that increased impervious area gives increased loads and concentrations for metals; increased impervious area gives increased loads, but decreased concentrations for nutrients. The greatly increased volumes of surface runoff from impervious (and directly connected areas) are largely associated with larger loads. Actual land use and surface characteristics within the urban area has been shown to result in differences in loading in some cases, particularly when commercial areas are frequently swept (Sartor and Boyd 1972; Bradford 1977). Similarly, in studies of different road surfaces, asphalt surfaces have been found to contribute higher loadings than concrete (Pitt 1979). The surface condition generally seems to be the most important factor in these studies, at least for solids washoff. As might be expected, surfaces with much free material sitting loose on the surface produce higher loads. The scatter in land use comparison data is large, however, and directly connected impervious area - rather than urban land use - typically has the strongest with pollutant loads. The scatter in data from different land uses is large and directly connected impervious area - rather than urban land use - typically has the strongest relationship with pollutant loads (Duncan 1995a). In addition urbanisation itself is responsible for increased levels of some pollutants that contribute to stormwater pollution: large urban areas (New York and London) have been shown to have greater concentrations of anthropogenically produced metals (Ca, Cd, Pb, Cr, Zn, Cu) than smaller urban areas (Halifax, Christchurch, and Kingston) in airborne dust (Fergusson and Ryan 1984).

2.2 The key stormwater pollutants

Analysis of French studies from the 1980s concluded that solids could be regarded as one of the basic indicators of urban runoff pollution, as many problem pollutants, metals and some nutrients are largely transported attached to particles (Desbordes and Servat 1984). Problematic pollutants, such as metals and attached nutrients, have been specifically associated with the fine fraction of TSS (Sartor and Boyd 1972; Deletic and Orr 2003). Prediction of TSS concentrations is therefore considered as a primary requirement for stormwater modelling and management. However there are also pollutants that are highly dissolved such as some nutrients (NO_3, PO_4) and metals (Zn, Cu) that have different methods of generation, and different impacts on receiving waters.

The concentrations of pollutants change with the flow, however most studies report on the ***event mean concentration, EMC***, defined as the event pollutant load divided by the total event runoff volume. It has been shown that EMCs are log normally distributed for most pollutants (Duncan 1999; Fuchs et al. 2004). Some studies report on ***site mean concentrations***, ***SMC***, defined as the total load divided by the total flow from a catchment (City Design 2003). This is a practical catchment characteristic that could easily be used to quickly assess annual pollution discharges from a given site or catchment.

Several studies have compiled worldwide data sets of urban runoff pollution, the most recent one being done by Fuchs et al (Table 2.1). Interestingly, more recent studies have shown that several important trends over time are observable in urban pollutant runoff. The compilation and updating of three major US programs undertaken from the early 1980s to the late 1990s showed significant changes in concentration of many parameters (Table 2.2). For example, concentrations of lead have fallen, due to the removal of lead additives in petrol. In several recent studies TSS concentrations have been reported as being smaller than in earlier studies. This is an important observation, since design of stormwater treatment systems rely on these 'general estimates' of pollutant loadings. It is possible that some of these differences over time may be due to changes in sampling methods (Smullen et al. 1999), but it is also possible that changes in land use, infrastructure and technology inherent within urban areas, are contributing to real changes in pollutant loading rates.

Table 2-1 Typical concentrations of worldwide stormwater pollutants (adapted from Fuchs et. al (2004))

Parameter	**Number of records**	**25th percentile**	**Median**	**75th percentile**
TSS	178	74 mg/l	141 mg/l	279.7 mg/l
BOD	88	8 mg/l	13 mg/l	20 mg/l
COD	136	51 mg/l	81 mg/l	113.3 mg/l
TOC	14	16.5 mg/l	19 mg/l	28 mg/l
TP	149	0.24 mg/l	0.42 mg/l	0.70 mg/l
NH_4-N	41	0.5 mg/l	0.8 mg/l	1.20 mg/l
NO_3-N	110	0.60 mg/l	0.80 mg/l	1.49 mg/l
TN	17	2.1 mg/l	2.36 mg/l	5.8 mg/l
Cd	54	1.16 µg/l	2.3 µg/l	5.0 µg/l
Pb	158	46 µg/l	118 µg/l	239 µg/l
Cu	127	28 µg/l	48 µg/l	110 µg/l
Zn	128	128 µg/l	275 µg/l	502 µg/l

Table 2-2 Comparison of compiled US data with NURP project data (adapted from Smullen et. al (1999))

	Parameter (Mean EMC)						
	TSS (mg/l)	**BOD (mg/l)**	**TP (mg/l)**	**Nitrite/ Nitrate (mg/l)**	**Copper (µg/l)**	**Lead (µg/l)**	**Zinc (µg/l)**
Compiled US Data (Smullen et	78.4	14.1	0.315	0.658	13.5	67.5	162

al 1999)							
US NURP Data	174	10.4	0.337	0.837	66.6	175	176

It has been hypothesised that catchment land use has an important impact on stormwater quality, but compilations of stormwater quality data have found that distinctions between typical urban catchments (i.e. commercial, industrial, residential) are minimal or not significant, in part because the variability in concentrations within a given land use tend to be as great as variations between land uses (Duncan 1999). At a small scale specific studies of roofs or roads have shown some differences (Yaziz et al. 1989; Clark 2006).

2.3 Description of processes

Descriptions of urban stormwater pollution are often divided into a number of processes, the end results of which are pollutant loads arriving at receiving water bodies. Several published reviews give descriptions of these processes (Bertrand-Krajewski et al. 1993; Duncan 1995; Zoppou 2001). The most commonly described are atmospheric deposition, buildup of pollutants on surfaces, washoff of those pollutants and transportation. These processes are not independent, as the amount of atmospheric deposition impacts the amount (Chui et al. 1982) and the composition of buildup (Fergusson and Ryan 1984). It is difficult to separate the transportation and the washoff of pollutants, particularly in separate stormwater systems where material in pipes is a result of washoff, as sampling is usually performed at the end of a pipe or stream section.

2.4 Buildup

A much referenced description of stormwater quality was that of Sartor and Boyd (1972), described in Sartor et al. (1974). The study looked at the principal constituents of material on urban surfaces, the difference in buildup on different surfaces and the effectiveness of management techniques at the time (mainly street sweeping). Buildup was found to be greater in industrial areas – attributed to poorer surfaces and less street sweeping – and largely inorganic. Most pollutants were associated with the finest fraction of the buildup material.

Buildup has been measured by a variety of methods such as dry vacuuming (Freund and Johnson 1980; Grottker 1987; Ball et al. 1998), brushing and vacuuming (Vaze and Chiew 2002), vacuuming and washing (Orr and Deletic 2003) and using industrial vacuum systems designed to remove the fine fraction (Butler et al. 1992).

Many studies have shown that buildup increases with time since last rainfall (Sartor and Boyd 1972; Pitt 1979; Grottker 1987; Haster and James 1994; Ball et al. 1998). Some studies conclude that buildup increases with time (Novotny et al. 1985) however other researchers do not find the increasing buildup relationship (Charbeneau and Barrett 1998; Orr and Deletic 2003). Studies repeatedly show that surface loadings of solids are comparatively large when compared to those reported in studies of washoff (Duncan 1995).

Descriptions of buildup have generally centred around the opposing processes of accumulation after rainfall, street sweeping and sometimes wind (Ball, Jenks et al. 1998), and the deposition of material from wind, traffic and other effects (Shivalingaiah and James 1984a). Large storm events may also deposit material from pervious onto impervious surfaces (Pitt 1979). Theory suggests that after some period the opposing processes will reach equilibrium, and any further material deposited will rapidly be redistributed. Willkinson (1954) as quoted in Weibel et al. (1964) found that the time to reach equilibrium was approximately 10 days. Grottker (1987) found that the accumulation was most rapid in the first 24 hours after cleaning. The scatter in the data is, however, quite large in most published studies.

While most studies have focussed on solids, buildup is likely to be different for different pollutants. Rainwater contains significant amounts of nitrogen (Malmquist 1978), and sometimes rainfall loads of nitrogen exceed those of runoff (Characklis et al. 1978). Clearly in this case the process of buildup and washoff is complicated by other factors; the buildup function will therefore need to be carefully assessed and applied for different pollutants and situations.

Mathematical descriptions of the process of accumulation have taken various forms, typically in their relationship with antecedent dry period. Sartor et al. (1974) presented exponential buildup curves for industrial, commercial and residential land use. Data from different catchments was lumped together and trimmed to form the accumulation curves. Even given this the data does not seem to support an exponential function above others.

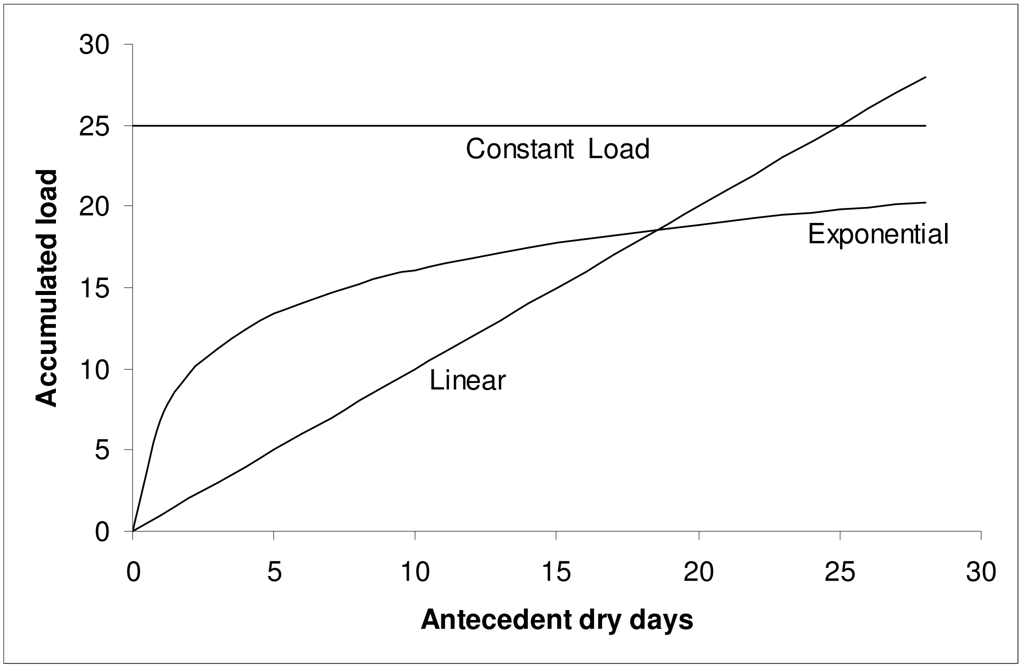


Figure 2-1 Simplified figure representing three different build-up functions with antecedent dry days

The function is presented in a form that suggests buildup starts near zero at the end of each washoff and street sweeping event. The authors noted several of these issues and limitations, however the exponential function presented has formed the basis of many descriptions since that time. Some researchers such as Alley and Smith (1981) have used a straight line to describe buildup, however they also state that an exponential accumulation curve is more general than the linear one used in many models (Figure 2-1). Ball et al. (1998) tested linear, exponential, power, reciprocal, and hyperbolic functions in a study in NSW and found that both street sweeping and rainfall reduced the surface loads. The power relationship provided the best fit, but as with most previous studies the scatter was large. Exponential buildup as implied by Sartor and Boyd (1972) remains the most common description used for detailed physical modelling.

Most authors acknowledge that buildup does not increase infinitely with antecedent dry period; at some point in time the competing process of distribution by wind (and other processes) will restrict the amount of material on the surface. Novotny et al. (1985) recommend that a decreasing rate of increase model is most appropriate. Haster et al. (1994) state that the underlying function is difficult to predict with accuracy and that site-specific data should be used. SWMM version 5 from the USA EPA (April 2008), offers users the choice of saturation, exponential, power or constant load functions. In recognition of the scatter in published data, other authors have proposed that antecedent dry period is a poor descriptor and have used a log-normally distributed random variable to describe initial load (Charbeneau and Barrett 1998).

Some authors have questioned the importance of buildup in predicting consequent washoff loads or concentrations. A modelling exercise by Smith (1997) used a simplified version of the well studied SWMM model and data from five catchments in Sydney and Canberra. A total of 35 events were used for TSS, 13 for TP and 12 for oxidised nitrogen. The study tested linear, constant surface load and exponential accumulation relationships. The linear function performed least well, as it is not bounded, and surface loads were often several orders of magnitude greater than the exponential and constant load functions. Predicted results from exponential and constant load methods were comparable. This suggests that the form of the buildup equation may not be a critical part of washoff modelling, at least for the storms tested. Similar results have also been demonstrated in later studies by Chiew et al. (1997).

Perhaps, most importantly, accumulation on the surface does not automatically translate to large loads in washoff. Malmquist (1978) repeatedly washed a catchment in Sweden with simulated rainfall equivalent to a relatively intense storm; four repetitions were required before washoff loads declined markedly. In an experimental study in Melbourne Vaze and Chiew (2003b) found that washoff loads were approximately 5% of total material on the surface. Studies in the 1980s in Washington State, USA after the eruption of Mt St Helens show that even with an excess of surface material, the washoff load is limited (Asplund et al. 1982; Chui et al. 1982) by other factors. In the Mt St Helens case even with several centimetres of surface loading the washoff load only doubled.

A study from Melbourne proposed that accumulation of material could be separated into two fractions, free and fixed loads (Vaze and Chiew 2002). The definition of the loads was a practical one; the road surface was first vacuumed (free load) and then scrubbed and vacuumed (fixed load). Although the study only covered three rainfall events, five drizzle events (no runoff) and five street sweeping occasions, it presents an interesting description of buildup. Free load decreased after each rainfall event. Total load decreased after two of the events, but marginally increased after the third event, which was a long (>24hrs), low intensity event with very little runoff. Consistency between daily accumulation measurements was low; therefore results are in no way conclusive. The results strongly suggest, however, that total buildup, even of the free, easily removable load, is not reduced to zero after small rainfall events. The suggested model has interesting implications for stormwater quality prediction. Particles found on impervious surfaces contain both easily removable and tightly attached varieties, and probably a continuum between the two. Significant energy is required to remove the attached particles; therefore only very long intense storms will begin to exhaust this reservoir. Freely moveable particles are replaced quickly through wind, traffic breaking up the attached load or street sweeping in-between rain events. The computer model MOSQITO from Hydraulics Research, Wallingford UK (as described in Bertrand-Krajewski et al. (1993)) also separates the accumulated fractions into non-cohesive and consolidated sediments. Results from this model, however, have not been better than comparable models where buildup is considered homogeneously, and it has not received widespread use.

Despite many studies the representation of buildup for modelling purposes is still an uncertain process. Early assertions that significant percentages of available pollutants were regularly removed from surfaces seem to have little support. The relationship between buildup loads and those in stormwater washoff is at best undefined, and the improvement in modelling of washoff loads by using buildup uncertain.

2.5 Washoff

Washoff is defined as the removal and transport of pollutants from surfaces. Physical descriptions of washoff, at least for particulate pollutants, centre on two physical processes; the removal of particles by either rainfall drop impact or by shear stress from overland flow. Researchers have proposed many measurable variables that represent these physical processes; runoff volume (Bedient et al. 1980; Freund and Johnson 1980), runoff rate (Alley et al. 1980; Aalderink et al. 1990), rainfall intensity, (Sartor and Boyd 1972; Price and Mance 1978; Jewell and Adrian 1982; Kuo et al. 1993; Yuan et al. 2001) or shear stress (Akan 1987). Some of these studies consider the underlying processes involved; others have been determined through statistical analysis. Often the washoff description is paired with some form of buildup function.

Many of these factors are related; for instance rainfall intensity is the main driver behind runoff rate from impervious surfaces. This makes separation of the influence of the individual variables difficult, and presents difficulties in designing stormwater-monitoring programs to accurately represent the different processes. It is also likely that different processes dominate in different situations. The link between shear stress in channels and sediment mobilisation is well studied in the geomorphologic literature (Hammer 1972).

Raindrop kinetic energy is often orders of magnitude greater than that supplied by overland flow (Hudson 1971). The subject of representing kinetic energy with rainfall intensity is well covered in studies of erosion of pervious areas with a review of various methods (van Dijk et al. 2002), suggesting that relationships vary with various climatic factors, but that rainfall intensity at short timesteps can be used as an effective surrogate indicator.

It is used to represent the kinetic energy of the falling raindrop and has been considered along with other descriptive variables such as the product of the drop size and velocity; and the product of the raindrop momentum and raindrop diameter (Salles et al. 2000). Other pervious area studies have shown that rainfall energy is of critical importance in the erosion of unprotected soils. Young and Wiersma (1973) used mesh screens to reduce the kinetic energy of rainfall and measured soil loss for different soil types. An 89% reduction in kinetic energy resulted in a 90% or more reduction in soil loss. It was concluded that rainfall kinetic energy, which may be represented by intensity, is a dominant process in the removal of particulates from pervious surfaces.

Jewell and Adrian (1982) used regression analysis to test parameters for catchments in the United States (US). Various catchment and storm characteristics were tested for both total event loads of different pollutants, and for instantaneous fluxes during events. The variables tested included preceding dry days, intensity of runoff (in m^3/hour), the volume of runoff, total rainfall, the percentage impervious area and storm duration. Rainfall intensity was discussed as a variable but not used in any of the equations tested. Although high R^2 values were obtained for most catchments, no one set of variables was consistently better than others at predicting pollutant washoff. Catchments from similar geographical areas that were modelled with the same model showed a large difference in calibration parameters.

Building on the concept of different processes for different pollutants, some studies have suggested that a combination of shear stress and rainfall intensity is the best predictor. Deletic et al. (1997) produced a model that includes particle detachment from raindrop energy and total shear stress. The proposed model simulated both flows and TSS loads (based on turbidity) on small (approximately 200-300 m^2) catchments. For 18 and 39 events in Belgrade and Lund respectively, the regression model produced R^2 values of 0.65 of predicted versus observed values in both catchments.

Irish et. al. (1998) refined this method specifically for highway runoff. Antecedent conditions, runoff intensity and runoff volume were found to be the most important variables in a study that also considered traffic loads, but did not specifically include rainfall intensity. This study also distinguishes between the causal variables for different pollutants. TSS is affected more by antecedent conditions than nitrogen or phosphorus, which the study concludes are largely contained in the actual rainfall rather than on the highway surfaces.

A study by Vaze and Chiew (2003b) repeated the work of Young and Wiersma (1973) for impervious areas. Known loads were placed on a small plot both with and without mesh screens to test the effect of reduced rainfall based kinetic energy. Loads from “events” generated by a rainfall simulator when screens were in place were less than half those without. Washoff pollution was still present in those experiments with mesh screens indicating that runoff processes, or very low rainfall intensity can contribute in some situation. Another interesting finding of this study was that all simulated storms, with average return intervals up to 1-2 years for Melbourne, removed less than 5% of the total material on the surface.

Associated with the washoff process is the description of variations in pollutant behaviour during stormflow – the first flush. Much has been written about the concept of "first flush" in the scientific literature including much debate about the definition of the pollutant behaviour (Bertrand-Krajewski et al. 1998; Deletic 1998). Generally the first flush refers to a situation where the pollutant cumulative load exceeds the cumulative flow early in an event; however some authors have used definitions involving concentrations (Odnevall-Wallinder 2000).

Definitions of first flush have varied widely; Geiger (1987) suggested the a first flush occurs when the slope of a normalised mass vs volume graph is greater than 45%. Bertrand-Krajewski et al. (1998) used a very strict definition, when 80% of the pollutant load is transferred with the first 20% of runoff.

Associated with the load behaviour is a pollutant concentration peak preceding the flow peak. Spangberg and Niemczynowicz (1993) compared the timing of the rainfall, flow and pollutograph peaks on a small experimental catchment in Lund, Sweden. The pollutograph peak consistently lagged behind the rainfall peak but was consistently before the flow peak.

Depending on the definition, first flush has been demonstrated to occur on small catchments (Yaziz, Gunting et al. 1989), and in some cases related to the antecedent dry period (Mason et al. 1999). In some small catchments studies have failed to detect first flush behaviour, at least for TSS (Deletic 1998). As catchment size (and time of concentration) increases first flush becomes less clear (Vorreiter and Hickey 1994; Chiew, Duncan et al. 1997; Cristina and Sansalone 2003; Cristina and Sansalone 2003). As time of concentration increases to equal then exceed storm duration the first flush from different parts of the catchment contributes at significantly different times to washoff, therefore negating any first flush effect (Cordery 1977).

2.6 Other processes

In addition to buildup and washoff there are a variety of factors that could affect either total runoff loads or the buildup and washoff functions themselves. For instance various studies have related pollutant loads and types to traffic volumes (Shaheen 1975; Reinertsen 1981; Shivalingaiah and James 1984b). In some cases it is almost certainly the main variable associated with pollutant mobilisation.

Catchment slope was described as one of the factors that affects washoff decay (Nakamura 1984). Slope has also been considered with respect to velocities when transport or overland flow is considered. It is likely that slope has several effects on washoff:

- Steeper slopes allow water to flow quickly from impervious surfaces, reducing any "cushioning" effect between impervious surfaces and raindrops
- Increasing velocities within channels and pipes, therefore keeping particles in suspension for longer
- Increasing shear stress associated with overland flow

Transformation of pollutants within the drainage system is another process that has been shown to have impacts, particularly in the combined drainage systems of Europe (Bertrand-Krajewski and Scrivener 1993). Specific activities such as construction activities (Bedient et al. 1980; Sharpin 1995), deliberate release of materials and septic tanks may also affect total loads in some cases. These processes tend to be isolated to specific cases, and operate in addition to a "regular" accumulation/washoff process. Their modelling, depending on the scale required, is most likely to require a stochastic approach.

All of these factors need to be considered when assessing pollution from drainage networks and its likely impacts.

2.7 Modelling Approaches

The consistent and accurate prediction of stormwater quality has proven to be an elusive goal. Reviews of available models have consistently stated that no model is best for every situation, complex models are difficult to verify and calibrate, and that even the most physically based models contain high levels of uncertainty (Huber and Heaney 1980; Bertrand-Krajewski et al. 1993; Ahyerre et al. 1998). Consequently there are a variety of approaches taken to predict the quality of urban runoff with the choice of approach being largely based on the reason for modelling.

Descriptions of the processes involved can be grouped into atmospheric deposition (wet and dry), interception by plants and buildings, buildup, washoff and transport (Duncan 1995); other authors have also considered transformation of pollutants within piped systems, particularly when considering combined sewer systems (Bertrand-Krajewski et al. 1993). Of course these processes are intimately related to the cycling of pollutants throughout the environment and in turn affected by the production of atmospheric pollution and the capacity of receiving waters to assimilate pollutants.

Many authors have divided models into groups based on the modelling approach (Huber 1992; Bertrand-Krajewski et al. 1993; Zoppou 2001). Models are generally grouped according to their level of complexity; physical representation attempted or planned use. It is acknowledged that the classifications used in this section are not completely distinct, but are presented largely for discussion purposes.

Statistical approaches

Statistical approaches generally use data from monitoring programs to relate hydrological or catchment characteristics to observed pollutant loads or concentrations. Event Mean Concentration (EMC) models statistically describe the long-term production of pollutants; the EMC is then typically multiplied by the event volume to determine the event load. Site Mean Concentration(SMC) models use a similar approach, although the mean concentration is derived from the total flow at a site, therefore providing a flow-weighted measure. Most EMC (or SMC) models assume log-normally distributed concentrations for urban stormwater (Huber 1992). This assumption has been shown to hold for Australian and international datasets (Duncan 1999; Brezonik and Stadelmann 2002) including the Blackburn Lake and Brisbane data discussed in chapter 4 (Francey et al. 2004).

Charbeneau and Barrett (1998) discuss the use of EMC models as part of a wider comparison of model types and conclude that the EMC method is adequate for longer-term impact assessment on receiving water bodies. One limitation of this approach is that no intra-event processes are considered and therefore the change in concentrations within an event cannot be described. The other limitation is that base flow related pollutants are not adequately represented in most cases, if at all.

The difference between EMCs, even for catchments with similar land use can be significant (Chiew and McMahon 1998), therefore the use of these models without calibration to a particular catchment involves significant uncertainties. EMC models can, however, produce good predictions of loads in the long to medium term when calibrated and therefore they are appropriate for uses such as the prediction of annual pollutant loads. However, a large number of models attempt to model the concentration and flow separately and then multiply these if loads are required (Huber and Dickinson 1988). Little discussion has been undertaken in the literature about what the fundamental variables are – loads or concentrations. Because concentrations are relatively easy to measure they have generally been used.

Probabilistic models (Rossi et. al. 2005) offer a more sophisticated version of this approach, by representing and selecting concentrations based on a predetermined distribution. The most commonly used stormwater quality model in Australia is MUSIC (CRCCH 2003), which uses stochastically generated concentrations based on EMCs at each model timestep. A separate pollutant distribution is included for base flow conditions. As with any EMC model coupled to a hydrological model, event loads from MUSIC are only based on physical catchment characteristics (such as area and land use), the total event runoff and measured EMCs.

Towards physical description: regression models

The French national stormwater quality monitoring program of the 1980s used a multiple regression approach to compare TSS concentrations and a variety of storm characteristics for four urban catchments; two in the Paris region and two in the south of France. Variables used were maximum flow rate, maximum rainfall intensity over a five minute period (Imax5), maximum intensity over the time of concentration, total rainfall depth, rainfall duration, runoff duration, preceding dry days (DTS), rainfall depth over the seven preceding days and rainfall depth since last runoff event. High regression R^2 values (~0.9) were obtained for mean concentrations and Imax5 in the south and moderate values (~0.6) for mean concentrations and DTS in the Paris catchments. Runoff volume was not presented in the multiple regression analysis, but it was used in further analysis (principal component analysis). The unexplained differences in the calibration factors inhibit the use of the equations derived beyond the monitored catchments. Desbordes and Servat (1987) produced two equations based on the data from the four catchments to formulate equations for TSS, BOD and COD loads. Imax5 and runoff volume were used as variables in the first model and surface loading was added in the second. Both models performed similarly, with errors of about 10% for annual totals and about 30% for individual events.

In the 1980s, the U.S. Nationwide Urban runoff program (NURP) measured stormwater runoff at approximately 81 sites in 28 cities across the USA. Regression analysis was used to identify the most important causal variables behind both loads and concentration generation (Tasker and Driver 1988; Driver and Troutman 1989). Storm variables tested included total rainfall, storm duration and maximum rain intensity in a 24 hour period. This study used the explanatory variables of total rainfall and catchment area to produce regression equations for three rainfall regions within the USA for both pollutant loads and runoff volumes. Total storm rainfall and total drainage area were found to be the most significant explanatory variables. Pollutant load results were significantly better predicted for dissolved pollutants than for TSS. Further work on these data sets suggested that the equations used could be improved; work done included a term for antecedent dry period which improved the regression equations for loads in the high rainfall areas but had little, or no effect in other areas (Driver 1990).

Importantly, Huber (1992), in a discussion of this approach, notes that when loads are calculated there is a spurious correlation between volume or flow related variables and the total load since load is the product of the flow and concentration. However, whilst this concept of 'spurious self-correlation' has been noted in the literature (Kenney 1982), there is also a genuine risk of ignoring useful relationships between one (easy to measure) variable and another (more difficult to measure) variable (Prairie and Bird 1989). Whilst pollutant load is a fundamental variable in its own right, flow and concentration are much easier variables to measure directly than are pollutant loads and attempts to relate the two are both understandable and potentially useful for development and implementation of stormwater management strategies.

Brezonik and Stadelmann (2002) presented another large data set (approximately 500 events over a number of sites) from the eastern US. Drainage area, total precipitation and average rainfall intensity (over the event) were used to produce regression equations for various pollutants, however only 33-57% of the variance in loads was explained by these models.

Duncan (1995), in a review of urban stormwater processes, notes the various methods of representing rainfall intensity in stormwater quality modelling and proposes an equation of the form:

Event Load = $\sum_{i=l}^{n} a(I)^{b}$ **Equation 2-1**

where *I* is the rainfall intensity over each time step of a model and a and b are calibration coefficients. The cumulative nature of this equation represents the ongoing input of energy through raindrop impact.

The model was tested by Vaze and Chiew (2003), using regression equations to compare the equation above, incorporating rainfall volume, runoff volume, runoff rate and a combination of the rate and intensity terms, with pollutant loads. A process-based model similar to the Stormwater Management Model (SWMM) (Huber and Dickinson 1988) was also used in the comparison. The catchments used were Blackburn Lake, Melbourne (discussed in Chapter 3); Sandy Ck, Brisbane; and Cressy St Brisbane a total of 20, 14 and 18 storm events were used respectively for this study. The results for the rainfall intensity, runoff rate and process based models were comparable; the results from the regression equation which combines rainfall intensity and runoff rate however were significantly better. This may be partially due to having two extra calibration coefficients, particularly when only a small number of events were being considered.

Brodie (2007) tested a very similar model to that in Equation 2.1 (based on the square of the rainfall intensity at 6-minute timesteps) on small (<450m^2) urban catchments and recorded only moderate agreement between the modelled and measured TSS loads. Somewhat similar approaches have also been tested outside Australia. Kanso et. al.(2004) modelled instantaneous concentrations of pollutants as $a \times I^b$ (where I is the instantaneous rainfall intensity, and a and b are calibration coefficients), and then calculated event loads. Using a limited data set they showed that this approach was inferior to the other three models used (they did though point out that this could be mainly due to the model having the least number of parameters of all the models tested). In the past, have many other attempts to use rainfall intensity as the key variable, using a number of very different (from Equation 1) approaches with mixed success, e.g. (Tasker and Driver 1988), (Desbordes and Servat 1987) (Huber and Dickinson 1988).

Physically based models

Many studies have gone further and attempted to use physical parameters and processes to improve prediction of pollutant loads/concentrations over short time steps. The variety of processes, many of which are related to each other, makes separation of the most important variables a complicated task. Buildup and washoff are usually seen as the most directly related processes to the pollutant loads delivered by urban runoff.

Washoff is often modelled using an exponential function (Sartor and Boyd 1972) of the basic form:

$$\frac{dP}{dt} = -kPR^{a}$$ **Equation 2-2**

where *P* is the remaining surface load, t is the timestep, *k* is a washoff co-efficient, *R* is runoff (or rainfall in some cases) raised to a power *a*. This equation, or derivations of it, are the basis of well-studied models such as SWMM (Bertrand-Krajewski et al. 1993). Initial models were proposed without the power *a*, however without the power the concentration cannot increase during an event regardless of conditions. Concentrations of some pollutants (such as TSS) do increase with increases in runoff rate, as shown in a study of over 200 storms in the Netherlands (Aalderink et al. 1990).

One of the main limitations of explanatory models such as equation 2.2 is that the coefficients k and a are largely empirical; k in particular can vary widely. Its value has spanned at least an order of magnitude (Ahyerre et al. 1998). Other authors have related k to catchment characteristics and total storm volume (Nakamura 1984; Charbeneau and Barrett 1998). The power term a has been shown to vary according to the type of pollutant being modelled (Sriananthakumar and Codner 1992), with values for suspended solids usually falling in the range of 1.4 –2.

Data used in the development of these models has often been a limitation, with models usually being tested using limited events or using averaged data – in particular rainfall.

Coleman (1993) compared SWMM with a “fundamental” model which had terms for particle detachment by rainfall impact and sediment transport with flow on two catchments in South Africa. The square of rainfall intensity was used to describe raindrop impact. No buildup term was included in the “fundamental” model as it was assumed that sufficient TSS was always available for washoff. Results from the two models were comparable with the fundamental model being easier to calibrate.

James and Shivalingaiah (1986) argue that given continuous modelling, the traditional buildup-washoff modelling approach should be replaced by a method that considers the dynamic equilibrium of accumulation and removal. A variety of the available models claim to supply satisfactory pollutographs for given rainfall events if enough data is available for local calibration, although many have been tested only on very limited data sets. However, the actual use of a pollutograph in an operational sense has also been questioned by several authors (Huber 1986). The level of temporal resolution required in pollutant load predictions will depend on the purpose of the modelling, and the nature of the stormwater treatment system being designed. For example, in a large system (with considerable storage), the temporal dynamics of flow may be important, but the temporal dynamics of water quality less so, since concentrations will be buffered within the storage.

Complexity in deterministic models when compared to simpler methods of prediction, limit their use for many of the needs described in *Chapter 1* (Shaheen 1975; Geiger 1984; Chiew et al. 1997; Smith 1997; Charbeneau and Barrett 1998). Direct comparisons of process-based models with simple EMC or regression analysis such as these, consistently show that the more complex approach does not significantly improve predictive performance (Geiger 1984; Vaze and Chiew 2003)), and yet increases the calibration data requirements.

2.8 Modelling uncertainties

A general observation from the literature is that physically based stormwater models are difficult to use and calibrate, and generally use largely empirical relationships. Models are an approximation of the real world, and therefore the processes they simulate are by definition incomplete. Models typically represent some major process (for instance rainfall in a streamflow model) and then use calibration parameters to simulate the complexities of real life (e.g. loss of water to evaporation or interflow). Their widespread application to highly variable urban catchments is therefore limited in practice. This is not to say that such models are of no use, however as Bertrand-Krajewski et al. (1993) described in a review of sewer modelling, the model and model complexity should be appropriately matched to the reasons for modelling in the first place.

There have been an increasing number of studies into the modelling uncertainties of stormwater models which have focussed on three sources of uncertainty; the uncertainty of the monitoring data used to assess model outputs, uncertainty which result from the model structure itself and the uncertainties in the input data used for the model. The data used to assess a model’s prediction has its own level of uncertainty which must be considered. In particular data used in stormwater management is often applied without detailed understanding of the inherent uncertainties associated with its measurement (Harmel 2006; McCarthy 2008).

Data uncertainties

In addition to the uncertainty inherent in the model structure there is uncertainty associated with the model input parameters, in the case of hydrological models the spatial and temporal variability of rainfall in particular (Jakeman and Hornberger 1993; Ogden and Julien 1993).

Uncertainties in total rainfall measurements are reported as 5 -15% at a particular gauge with uncertainty rising with increased rainfall intensity (Strangeways 2004). These uncertainties are related both to the placement of the rain gauge (i.e. raised, clear of obstructions) and the size of a tip. As the bucket on the gauge is only of a certain volume this will affect the quality of data at various rainfall intensities. If the bucket is very small, in high intensity events it will spend a lot of time "tipping" and therefore could under represent the true rainfall. The converse of this is a large tip size which will not tip as often and therefore will measure the total rainfall more accurately, but will not capture as much information about intensity.

Spatial variability also influences rainfall measurement. In even small urban catchments the spatial variability of rainfall has been shown to significantly affect the total volumes and the intensity reported within an event (Berndtsson and Niemczynowicz 1988; Nelen et al. 1992; Chaubey et al. 1999; Moreira et al. 2006).

In terms of discrete measurements of pollutants, some parameters show variance within the water column. Total Suspended Solid (TSS) measurements have been reported as having uncertainties from 15 -20% (Ahyerre et al., 1998; Bertrand-Krajewski et al., 2002, Harmel et al., 2006). Other pollutants with finer or more dissolved constituents (e.g. Total Nitrogen TN) have minimal associated sampling uncertainties as a result of better mixing in stormwater flows (e.g. 0% for TN – cited in Harmel, 2006). Analytical process are also a source of uncertainty. Figures of 10% have been reported for TSS with higher values expected for other constituents (Bertrand -Krajewski et al., 2002).

Two additional sources of uncertainty impact results when event loads or EMCs are calculated; the uncertainty in flow measurements and the method in which samples are combined to calculate an EMC. Flow uncertainties of 2 -20 % in using velocity-area methods have been reported in literature depending heavily on the equipment and the flow conditions (Harmel 2002). Samples can be taken on a time or a flow based basis, with flow-based sampling at short intervals giving lowest uncertainties (Harmel 2006).

The combination of these various sources of uncertainty is usually undertaken using the law of propagation of uncertainties (Talylor 1994). Studies in Europe have recommended a figure of 30% experimental uncertainty for long term TSS loads (Bertrand-Krajewski et al. 2002).

Model structure

Only recently has analysis of the relative importance of the model structure itself has been applied to typical stormwater quality models. Grayson et. al. (1992) noted that the model needs to reflect the complexity of the questions that are being asked. Over-paramertisation of models is acknowledged to not only introduce unwanted complexity, but also to hinder the use of models in other applications or modelling disciplines (Beck 1987; Ginzburg and Jensen 2004).

In a study testing the parameterisation of a commonly used stormwater model (SWMM) Gaume et al. (1998) performed an analysis and concluded that data availability in hydrology supports only models of limited complexity. Kanso et al. (2003) tested the sensitivity of parameters within four stormwater models and showed that various parameter sets could be generated that, when calibrated, find only local solutions and therefore "miss" the global optimum solution (Kanso et al. 2003)

These studies reinforce the link between the model structure, the data used to populate the model and the confidence in results. Further research is required to both assess the uncertainty associated with various modelling approaches and to assess this against the needs of the end users of models.

2.9 Summary of research

Regression approaches have often produced strong relationships between several variables and stormwater loads, although the reliable transference of equations between different catchments has not been widely demonstrated. Deterministic models have achieved similar results for a given catchment, but also have not been successfully applied between catchments without significant local calibration data. Deterministic models are also time-consuming to use and calibrate, and are perhaps better seen as a research tool rather than a management tool.

The often-stated advantage of deterministic or physical models is the ability to produce *pollutographs* – time series of pollutant flow rates or concentrations. Even if these were demonstrably accurate, the advantages of this approach are probably not extensive in stormwater management practice (Huber 1992) as treatment measure design or ecological systems are rarely sensitive enough to make very short term prediction necessary. The modelling of buildup and depletion over an event in particular has not been consistently shown to increase the accuracy of predictions. Accurate prediction at a short timescale however would indicate that the modelling approach is representing the physical processes involved rather that just statistically narrowing in on a mean EMC.

Various parameters have been used to describe pollutant generation, particularly for TSS, although consensus on the principal drivers is limited. Most descriptions rely on catchment characteristics and one, or a combination, of the parameters below:

- Total volume - therefore relating the pollutant load to the event rainfall
- Rainfall Intensity – which proposes that the energy supplied by rainfall influences pollutant removal
- Runoff rate – which can represent the shear stress supplied by overland or channel flow, but is also related to the intensity and duration of the rainfall

All of these factors are related for a particular storm event, and it is likely that with particular situations or pollutants that they all can provide satisfactory predictions. For instance a catchment that contains an eroding channel with high velocities of flow is likely to be influenced more by shear stress than other processes. Conversely an aluminium roof is less likely to be influenced by shear stress as the flow is distributed more evenly over the surface.

The erosive effects of increasing rainfall intensity, more particularly the power of the rainfall intensity, are accepted in studies of pervious area runoff as being important for predicting pollutant loads. Many studies have considered rainfall intensity to predict pollutant loads from impervious areas, but most have used an average intensity over the event, a maximum intensity or a long timestep. The French national and US NURP programs are arguably the two most influential monitoring programs in the field. Both considered rainfall intensity, although the approaches (maximum intensity in a 5 minute period or the maximum in a 24 hour period) achieved only modest results. When rainfall intensity has been considered over short timesteps it has usually been with limited data for comparison. Short-duration intensity, in its applications to date (Kanso et al. 2004; Brodie 2007), has thus shown promise, but without adequate data to thoroughly evaluate this promise.

These varied approaches, together with the uncertainty surrounding the temporal and spatial variability of rainfall, have resulted in no consistent method of using rainfall intensity for predicting pollutant loads. The approach of Duncan (1995) using a rainfall function, tested by Vaze and Chiew (2003), is suggested as a compromise between the deterministic approach and a simple management model. It acknowledges the central process behind pollutant generation but restricts the models complexity to that justified by the data available.

2.10 Statement of the problem

Current stormwater modelling methods are either statistically based or are complex in application. They also contain large uncertainties, particularly at the small time scales typically needed for stormwater treatment design and research needs. Often the complexity of the model is not reflective of the data available for operation. The number of variables involved and the stochastic nature of urban hydrology combine to present a complex situation.

The current study aims to investigate the primary variables which influence pollutant loads and, if applicable, use this to develop a modelling approach that has good predictive power, is transferable between catchments, user friendly and has known uncertainties.

The main objectives are:

1. To review existing relevant data sets on stormwater quality and conduct a large scale monitoring program of stormwater in Melbourne
2. Develop a simple predictive model for stormwater pollutants
3. Test the simple model for prediction of loads and concentrations at short timesteps
4. Complete sensitivity analyses of model parameters.

2.11 Key hypothesis

Many studies have found that rainfall intensity or runoff rate are good explanatory variables for pollutant washoff loads. Initial work by Duncan (personal communication) and a modelling exercise by Vaze and Chiew (2003) tested the proposed model on data from catchments in Brisbane and Melbourne. Vaze and Chiew also tested runoff rate in place of intensity. Results comparable or better than a typical build/up washoff model were reported. Further analysis undertaken as part of this study is presented and discussed below.

Although some authors have trialled similar approaches (Kanso et al. 2004), and the method has shown promise in desktop studies (Vaze and Chiew 2002b), studies have usually lacked the high quality data necessary to properly evaluate the method. This study will trial a number of datasets across various land uses and sizes, and will compare the uncertainty of the approach with the monitoring uncertainty. The primary data will also be taken in a relatively consistent climatic region, although several catchments with different rainfall patterns will be trialled initially.

Therefore the main hypotheses that will be tested are:

- That rainfall intensity is the primary driver for urban stormwater pollutant loads
- Pollution build up is not an important process in assessing loads of sediment and nutrients
- That high quality data is necessary to reduce modelling uncertainties to a level appropriate for practitioners

To test these hypotheses, it was deemed necessary to collect a large new data set. Therefore a group of secondary hypothesis is also to be tested, based on the collection of these data:

- Site mean concentrations of stormwater in Melbourne are of similar magnitude as worldwide means
- Catchment characteristics do not provide significant explanation of pollutant levels, relative to the variable proposed in Equation 2 -3

3 MONITORING PROGRAMS

3.1 Introduction

Catchment characteristics, landuse, size, and corresponding hydrological conditions are the most important factors in both the quantity and quality of the resulting washoff (Driver 1990; Charbeneau and Barrett 1998). Any analysis of pollutant generation needs to take into account a wide range of these factors to ensure widespread applicability. The hypotheses presented in Chapter 2 propose rainfall intensity as the main driver behind pollutant load generation, with catchment area and land use affecting the volume of runoff and hence the load of pollutant transported. To test this hypothesis a variety of hydrological conditions, in addition to a variety of land uses and areas, needed to be tested and compared. This chapter outlines the data used to test the hypotheses outlined in Chapter 2.

Data for undertaking the study were collected in two-steps:

- First, a review of existing data sets from both Australia and internationally was undertaken. A comparison of the monitoring methods, intents, and results was undertaken to determine their suitability for this study. Rules were then put in place to identify the data that could and could not be used, and to assess the quality of the data
- Second, an extensive and ambitious monitoring program, based in Melbourne, Australia was designed, arising out of lessons from the review of existing data sources, to test the main parameters responsible for washoff loads from a variety of catchments

This Melbourne monitoring program had a secondary aim of providing a local data set for characterisation of stormwater runoff. The data from this data set were used both to calibrate existing models and to provide a reliable, high quality data set for constituents, such as heavy metals, that are not regularly monitored in Victoria.

To provide a structure that recognises the two discrete sources of information, this chapter is divided into two sections:

- Part A consists of a review of existing data sets. The data sets were examined to ascertain both the suitability of the data for testing the hypothesis in chapter 2, and also to inform the scope of any new monitoring program necessary to provide a representative sample of urban pollutant generation
- Part B describes the methodology of the monitoring program set up specifically to test the hypothesis described in chapter 2 and provides an overview of the data set obtained – termed the primary data set. The methods, results, discussion, and conclusions are presented separately for each of the two sections. For example, the methods applied to review the existing data sets are described, followed by the resulting data set that was used for the analysis. The suitability of those data is then discussed. Similarly, the methods used in design of the Melbourne monitoring program are described, and the consequent data set presented

3.2 Secondary data sources

Ten data sets, containing data from 29 urban catchments, were examined for their suitability (see section 3.3 and Table 3-8). Most were from the Melbourne metropolitan region; however, two European catchments, one New Zealand and five Queensland catchments, were also tested.

Two primary tasks were undertaken in review:

- Development of a set of consistent rules that could be used to evaluate data from different sites and monitoring approaches. These rules could also be important because they might identify lessons learned from the existing data sets and could be used in the development of the specific Melbourne monitoring program. The rules were designed to ensure consistency and relevance of the data sets with respect to their suitability for testing the proposed pollutant generation model

- Statistical analysis of the data sets to determine the scope of monitoring necessary to describe pollutant generation to an acceptable level of uncertainty. Specifically, analysis was undertaken to identify the number of events needed to quantify concentrations of each pollutant within a level of uncertainty consistent with the capability of sampling and analytical methods

Rules for evaluation of existing data

All the data sets described were examined to determine the consistency of the data, the relationship between rainfall and runoff at the catchment scale, and the suitability for evaluation of the central hypothesis of this thesis. A set of rules were developed as a method of comparing results from different catchments and monitoring methodologies.

Two initial tests were applied in order to define what constitutes an acceptable storm event. They were:

- the number of samples taken to determine event load to determine if it was adequate to represent the event load
- the runoff coefficient (i.e. on rainfall/runoff behaviour)

There were also some other simple rules used in event definition, as discussed below. Finally, the data sets were examined to assess how the number of monitored events may impact on the characterisation of pollution loads.

Definition of an event based on the number of samples recorded

It is important to ensure that events used are based on an adequate representation of the event load (and its temporal variability). Pollutant concentrations are more variable in the ascending peak of the hydrograph (rising limb - when rainfall is occurring) and are more consistent on the descending peak (falling limb), (Lee et al. 2002). Therefore flow-weighted sampling that adequately captures the ascending peak is recommended (Leecaster et al. 2002). Sampling regimes that are time-weighted without regard to flow risk missing valuable information. Obviously it is possible to "structure" a time-weighted regime to sample more frequently at the beginning of an event, and indeed the Blackburn Lake program used this approach. This compromise is used on occasion by researchers because of limitations with auto sampler reliability under flow weighted programs.

Several sites had a very small number of samples describing the pollutant loads. There were also events with gaps of many hours between samples. The Blackburn Lake, Lund and Belgrade data sets were used to define an acceptable sampling regime as they had the most thorough coverage of pollutant samples over each event.

A volume-based approach to sampling was proposed, with a minimum of five samples proposed to define an event.

The other requirement was to define the maximum time between samples so that widely spaced samples, that may have missed rainfall peaks, were not considered representative of the event. This time was set at two hours, approximately four times the duration of concentration of the largest catchment in question. When two hours had elapsed without a sample being taken a new event was started. This, in effect, provided a method of separating events when the effects of rainfall (and therefore flow) were minimised.

Definition of an event based on rainfall/runoff behaviour

It is important that only “real” rainfall events be used, rather than events that occur due to some other source of discharge or equipment malfunction. In order to do this, only events with a “reasonable” rainfall-runoff coefficient ($r_c = \frac{v}{r \times A}$ Equation 3-1) were considered:

$$r_c = \frac{v}{r \times A}$$ **Equation 3-1**

where r_c is the runoff coefficient, r is the rainfall in metres, A is the catchment area (m^2) and v is the total runoff volume (m^3). A runoff coefficient of greater than 1 would indicate that there is more flow than rain recorded (or that baseflow is a major contributor to the flow volume).

According to Equation 3-1, an event was examined for extreme values compared to the “average” for the catchment. An extreme runoff coefficient probably represents a malfunction in the monitoring equipment or a rainfall event with high spatial variability – for instance a thunderstorm that affected only part of the total catchment (Sriananthakumar and Codner 1992a). An event with a runoff coefficient greater than 1 was rejected from the analysis.

Other rules for definition of an event

Consistent rules were also required to determine the start and end points of each event. Initial analysis involved the approximate calculation of the time of concentration (t_c) of each catchment (based on a velocity of 1 m/s over the longest travel distance). As shown in Figure 3.1, an event was defined as the rainfall that falls up to a time of initial t_c before the first sample was taken to the initial t_c before the last sample. This was performed for consistency across all catchments. The complete set of rules to define an event are listed in Table 3-1.

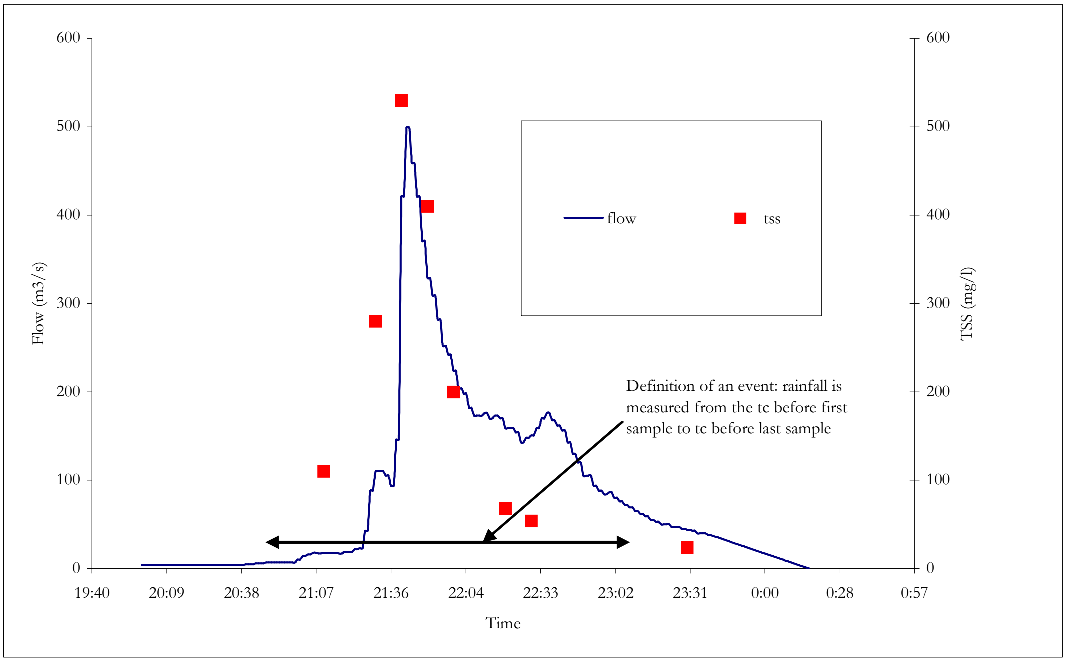


Figure 3-1 Definition of an event based on consideration of rainfall within, before and after the collection of samples

Table 3-1 Summary of rules used to define an "event"

Component	Rules to define an "event"
Rainfall	Taken from the time of concentration prior to a sample being taken to the time of concentration before the last sample. Must be greater than 0.2mm
Runoff	Taken from the time of concentration prior to a sample being taken to the time of concentration before the last sample. Must be greater than 0.2mm
Time between events	No rainfall for four hours. (This is approximately eight times the time of concentration of the largest catchment.)
Number of samples	Greater than five, with a maximum gap between samples of two hours.
Runoff coefficient (runoff divided by rainfall x catchment area)	Must be less than one. Also checked for consistency across time at each catchment.

Assessment of the variability of existing data sets with respect to number of events sampled

Given that one of the aims of the review of existing data sets was to inform the design of the primary data collection (see next section), analysis was undertaken using these existing data sets to determine the number of events necessary to categorise a particular catchment (in terms of event mean concentrations, EMCs) within a specified level of uncertainty.

The raw data sets available have been reported in the published literature without any data manipulation or evaluation. Therefore the errors reported are likely to be at the high end of a possible range (i.e., appropriate data-screening and validation methods would reduce the uncertainty of the resulting data).

Using observations from previous studies on this topic, two assumptions were made:

- Relative uncertainties of at least 30% are to be expected in monitoring of flow (Ahyerre et al. 1998). Seeking greater accuracy therefore is a case of diminishing returns as the monitoring program will restrict the ultimate uncertainty. Therefore the monitoring program should aim at characterising the behaviour to within approximately 30% of the mean value. Most values reported are for TSS. However the 30% uncertainty has been used as a guide for TP and TN on the basis that the majority of errors are in the flow and sampling methodology (Bertrand-Krajewski and Bardin 2002b), and are therefore not pollutant dependent.
- That the distribution of event mean concentrations (EMCs) are often log-normally distributed (Duncan 1999). Therefore the data sets were log-transformed and tested for normality before calculation of confidence intervals.

If enough events could be sampled to reduce the uncertainty to approximately 30% there would be some confidence that the monitoring program was providing enough information to describe the EMC to a level of uncertainty dictated by the monitoring procedure.

In order to evaluate the uncertainty of the existing data sets with respect to the number of samples taken, the data were first evaluated for normality (to see whether parametric approaches, which rely on the assumption of normality, could be used). The Shapiro-Wilk test (Shapiro and Wilk 1965) was used to test the sample of events against a normal distribution at the 0.05 level of significance. This test calculates a statistic (W) that assesses whether a random sample, $x_1, x_2, ..., x_n$ comes from a normal distribution . Small values of W are evidence of departure from normality.

The test statistic is:

$$W = \frac{(\sum_{i=1}^{n} w_i X'_i)^2}{\sum_{i=1}^{n} (X_i - X_u)^2}$$ **Equation 3-2**

where W is the test statistic, X’ are the ordered data, and X_u being the sample mean. W_1 is defined as:

$$W_i = MV^{-1}\left[(M^1V^{-1})(V^{-1}M)\right]^{-1/2}$$ **Equation 3-3**

where M denotes the expected values of standard normal order statistics for a sample of size n and V is the corresponding covariance matrix. Low values of W represent a departure from normality.

If a satisfactory approximation to a normal distribution was achieved (i.e., $p>0.05$) the analysis of uncertainty relative to the number of events sampled would then be carried out using the central limit theorem. The central limit theorem allows a known confidence interval to be determined given the standard deviation and the number of events in a sample. Log-transformation enabled the assumption of normality to be satisfied in eight of the ten cases (significance level $p > 0.05$ see Table 3.10), although the transformation improved the statistic in all cases. The log-transformed data sets were therefore used to produce curves showing the 95th percentile range (as a percentage of the mean) versus the numbers of sampled events for TSS, TP, and TN. The number of events needed to produce a 95th percentile range of approximately 30% was then read from the graph.

The 95% confidence interval was calculated using the area under a normal curve corresponding to z =1.96 in $z = \frac{x}{\sigma/\sqrt{n}}$ **Equation 3-4**

$$z = \frac{x}{\sigma/\sqrt{n}}$$ **Equation 3-4**

where σ is the standard deviation of the original data set (of *N*), *x* is size of the confidence interval for given *z*, and *n* is the sub-sample size (as above).

The *n* term was varied (five, ten, 15 etc. events) to show the effect that varying the sample size had on the confidence interval. In this way a confidence interval corresponding to an expected uncertainty generated by the monitoring procedure could be calculated.

3.3 Description of the secondary datasets

Having outlined the methods used to evaluate the quality and suitability of the existing data sets, this section presents a summary of the available data after the screening rules were applied. It explains the characteristics of each monitoring site, the catchment characteristics, and the number of events available for each pollutant type.

The Lund and Belgrade catchments captured surface runoff whereas at Blackburn Lake and Sandy Creek samples were collected from channel sections.

Blackburn Lake

The Blackburn Lake monitoring program was conducted over two years (1996/97) specifically to investigate the behaviour of Blackburn Lake, a constructed lake in eastern Melbourne (RossRakesh et al. 1999). Several subcatchments upstream of the lake were monitored as part of the study, which also included in-lake sampling. The largest of these subcatchments (C) is 221 ha in area of mixed residential and commercial land use. The drainage system of subcatchment C is largely an underground piped system, with a short section of open “natural” channel described below. The outlet of the subcatchment is upstream of any lake backwater effects.

Two rain gauges were available with rainfall readings available 6-minute intervals (defined as the rainfall timestep) close to the catchment: Mitcham Reservoir at the eastern (upstream) limit of the catchment, and Masons Road retarding basin, 1.6 km south west of the catchment (see lower left corner of Figure 3.2). The Mitcham gauge was used whenever data were available as it was within the catchment boundary. During events when the Mitcham gauge was not functioning correctly the Masons Road gauge was used as a replacement. Both rainfall gauges recorded rainfall in 0.2mm tips. It should be noted, however, that the timesteps between the raw data and the final report have been adjusted. Therefore rainfall is displayed in 0.1mm increments as a tip was distributed between two time steps (RossRakesh, personal communication).

Parameters monitored as part of the program were TSS, TP, and TN (discrete time based sampling). Turbidity, conductivity ,and temperature were also recorded at two-minute intervals.

The catchment is largely piped. However, it contains a small section (approximately 500m) of “natural” channel, as shown in Figure 3.2. Work was undertaken by Melbourne Water to prevent erosion in this channel both prior to, and in, the years following the study. There was also a significant subdivision under construction adjacent to the channel during 1996/97. These factors may have influenced pollutant generation from the catchment when compared to an established urban catchment. Further details about the catchment and monitoring program are described in RossRakesh et al. (1999).

Figure 3-2 Blackburn Lake Catchment showing locations of rain gauges

Fifty-five events spanning almost two years were reported as part of the Blackburn Lake data set. Duncan (2003, personal communication) explored the relationship between rainfall intensity and pollutant loads using the Blackburn Lake data set. This data set has also been used by Vaze and Chiew (2003) in a study that compared the relevance of rainfall, runoff volume, runoff rate, and rainfall intensity as parameters responsible for pollutant loads.

When analysed by Vaze and Chiew (2003), pollutant concentrations from the Blackburn Lake program were assumed to be constant for five minutes either side of a measurement. Events were then discarded when the pollutant sampling did not represent 75% of the total event duration. This method has the effect of rejecting most of the longer rainfall events as the auto sampler was set to time-based sampling, and would thus take all samples well before the event had completely finished.

In the analysis all 55 events reported in the original report (RossRakesh et al. 1999) have been considered. Several factors reduced the number of events used for analysis. In particular the runoff coefficient for individual events showed inconsistencies during the monitoring period. There is a significant trend of decreasing runoff coefficient in the 1997 events when compared to 1996 (average of 0.25 compared to 0.38). This trend is apparent despite a large range in the intensity and size of events. Events were checked both with and without baseflow being taken into account. However the trend was still apparent. The systematic drop in runoff coefficients across a wide range of events suggests a possible error in recording the flow. The ten events from January 1997 were therefore omitted from the analysis. A total of 36 events were consistent with the criteria defined in Section 3.2.

Table 3-2 Data summary, Blackburn Lake

Data Summary - Blackburn Lake	
Timestep for rainfall gauge	6 minutes
Rain gauge measurement	0.2 mm
Distance to rainfall gauge (from catchment centroid)	Mitcham Reservoir 1.4 km (MW gauge ref 229633) Masons Road RB 1.6km (MW gauge ref 229636)
Timestep for flow measurements	2 minutes
Parameters sampled	TSS, TP, TN
Method of pollutant sampling	Time based sampling, all discrete samples
Catchment size	221 ha
Total fraction impervious	0.58
Land use	Mixed urban use: (13% Commercial, 28% Industrial, 6% Parks/Public Use, 48% Residential and 5% Major roads)

Lund (Sweden) and Miljakovac, Belgrade (Serbia)

Two European data sets were obtained that contained rainfall intensity and TSS (derived from turbidity). The catchments were both small (less than 300m^2), and each of them drained to a single point. The Belgrade catchment received moderate traffic loads, and sediment was observed to accumulate along the curbs. No such effect was observed at Lund (Deletic 1998). The average slope at Lund was 2.3% and 1.2% at Belgrade (Belgrade also had a small pervious bank of 15.4% slope that was only engaged at high intensities). A detailed description of the sites and monitoring procedure is given in several other published references (Spangberg and Niemczynowicz 1993; Deletic 1998; Calabro 2001).

The studies were designed to investigate the behaviour of pollutants at a 10 second timestep. Therefore turbidity (which can be logged continuously) was used as a surrogate for TSS. A data logger was used to collect readings every 10 seconds. The turbidity readings were calibrated to TSS in a laboratory for the site and exhibited high correlations. The calibration was carried out for a series of events in the case of Lund, but for only one event at Belgrade. The relationship is stated as being stronger when TSS concentrations are > 50mg/l. At these concentrations the error is estimated at 10 –15% (Deletic 1998). When concentrations were below this, the turbidity meter was regarded as being less reliable, indicating greater uncertainties.

A model presented in Deletic (1998) predicted both flow and TSS loads. The pollutant generation model used antecedent conditions, rainfall intensity, and the shear stress associated with flow as causal parameters. When calibrated, the model achieved average coefficient of determination values of 0.65 between these inputs and the pollutant loads (Deletic, personal communication).

Table 3-3 Data summary, Lund

Data Summary - Lund	
Timestep for rainfall gauge	10 seconds
Rain gauge measurement volume	0.035 mm
Distance to rainfall gauge (from catchment centroid)	270m
Timestep for flow measurements	10 seconds
Parameters sampled	TSS
Method of pollutant sampling	Turbidity sensor calibrated to TSS
Catchment size	0.027 ha
Total fraction impervious	1.0
Land use	Carpark

Table 3-4 Data summary, Belgrade

Data Summary - Belgrade	
Timestep for rainfall gauge	10 seconds
Rain gauge measurement volume	0.25 mm
Distance to rainfall gauge (from catchment centroid)	20 m
Timestep for flow measurements	10 seconds
Parameters sampled	TSS
Method of pollutant sampling	Turbidity sensor calibrated to TSS (using only one event)
Catchment size	0.021 ha
Total fraction impervious	1.0
Land Use	Asphalt street

These studies are important due to the short timesteps used for flow/pollutant monitoring and the fine scale of the rainfall monitoring. The use of turbidity as a surrogate for TSS, however, adds a source of uncertainty to the procedure, particularly in the case of Belgrade where only one event was used for calibration.

All events available (68 at Belgrade and 23 at Lund) were used for analysis as extensive investigation had occurred previously into these data sets, and therefore all events were found to satisfy the evaluation criteria described previously, in Section 3.2.

Sandy Creek and Cressy St, Brisbane
Data from Brisbane City Council's stormwater monitoring program were obtained which include data already discussed in several publications including attempts to relate rainfall intensity to pollutant loads (Walden 1999; Vaze and Chiew 2003). The data were originally collected primarily for management purposes.

The Sandy Creek catchment is located in the inner south-western suburbs of Brisbane (City Design 2003). Land use is mainly residential, with some commercial development and areas of open space. Topography ranges from moderately steep (5 – 10%) upstream to flat near the sampling site. The drainage system is mainly piped, with an open channel toward the downstream end. Brisbane has a sub-tropical climate with an annual rainfall of 1200mm, most of which falls over the summer period.

The Cressy Street catchment is located approximately 8km north of Brisbane's centre (City Design 2003). Land use is mainly residential. The catchment is relatively flat, with a maximum slope of approximately 2%. The drainage system is a separate, piped system. Site descriptions and monitoring methods have been described by City Design (2003). All sites were equipped with flow and pollutant samplers that recorded flow at one minute intervals, and collected pollutant samples using a time based interval regime. Rainfall was collected at a one-minute timestep; however, the data available only had 1mm tip sizes.

All events published by City Design (2003) from both sites (94 for Sandy Creek and 104 for Cressy Street) were considered for analysis. Many of the events consisted of widely spaced pollutant samples, small numbers of samples per event, and sampling that did not match rainfall records. Most therefore did not fit within the rules describing an event (Table 3-1) with only 32 events being considered suitable at Sandy Creek (for TSS) and only four at Cressy Street. The Cressy Street site was therefore not considered for further analysis.

Table 3-5 Data summary, Sandy Creek

Data Summary – Sandy Ck	
Timestep for rainfall gauge	1 minute
Rain gauge measurement volume	1 mm
Timestep for flow measurements	1 minute
Parameters sampled	TSS, TP, TN
Method of pollutant sampling	Discrete sampling
Catchment size (ha)	227
Total fraction impervious	0.42
Land use	Primarily residential

Table 3-6 Data summary, Cressy Street

Data Summary – Cressy St	
Timestep for rainfall gauge	1 minute
Rain gauge measurement volume	1 mm
Timestep for flow measurements	1 minute
Parameters sampled	TSS, TP, TN
Method of pollutant sampling	Discrete sampling
Catchment size (ha)	107 ha
Total fraction impervious	0.44
Land use	Primarily residential

Number of samples available

Table 3-7provides a summary of the number of samples available, for each pollutant, at each site. Blackburn Lake provides perhaps the best data set, with 36 samples available for each of TSS, TP, and TN. The Lund site is also a valuable resource, with 68 events monitored for TSS alone (based on calibration to recorded turbidity).

Table 3-7 Number of events considered for analysis after evaluation

Catchment	Pollutant	Number of events collected	Comments
Blackburn Lake	TSS TP TN	36 36 36	As discussed in Section 3.3, events in the later part of the monitoring program showed an unexplained decline in runoff coefficients.
Sandy Creek	TSS TP TN	32 17 17	Many events were discarded due to the small number of pollutant samples over the event.
Lund	TSS	68	TSS was derived from calibrated turbidity measurements.
Belgrade	TSS	23	TSS was derived from calibrated turbidity measurements (calibrated using one event).

3.4 Consideration of other monitoring programs

A number of other existing data sets were investigated to determine their suitability for the study. Table 3-8 shows the data sets considered but not used, the reason for their exclusion, and their published reference.

Table 3-8 Miscellaneous data sets and reasons for exclusion

Data set	Reason for exclusion	Reference
GHD, Melbourne (13 catchments across Melbourne)	Few events recorded at each site with few samples taken during each event. Rainfall gauges also not close to some sites.	(Gutteridge Haskins and Davey and Environment Protection Authority of Victoria 1981)
Coburg, Victoria	Data set lost	(Allison et al. 1997)
Muthukaruppan (4 catchments across Melbourne)	Maximum of 5 events at an individual site, some sites had no nearby rainfall gauges	(Muthukaruppan et al. 2002)
North Shore City Council, Auckland, New Zealand	Only 4 events available	(North Shore City Council 2002)
Brisbane City Council (Brisbane Forest Park, Wynnum, Keating Street, Cressy Street)	Mostly composite sampling (except Cressy Street). The timing of samples was not recorded in most cases so there was difficulty in matching rainfall to a particular sample time. The Cressy Street data were examined in detail; however very few events had a runoff coefficient less than 1. Often events also had less than 4 pollutant samples.	(City Design 2003)
Lynbrook estate	Only 10 events. No nearby rain gauges (nearest gauge was 3.5km away).	(Lloyd 2004)
Aderdeen, Scotland	Few events captured, with few samples taken for each event.	(Orr 2002)

3.5 Secondary data sources - Analysis

Variability of data with respect to number of events sampled

The available data was assessed for two reasons to determine level of variation in EMCs across the secondary monitoring sites and to inform the design of a future monitoring program. Data sets were taken before data evaluation, and were chosen solely on the number of events available (Table 3-9). The Belgrade data set was excluded due to the small number of events recorded there (23). This was a deliberately conservative approach which requires acknowledgement of a less than perfect (or perhaps typical) monitoring program design into the analysis. While performing the analysis data that had been refined would give a different result (smaller number of events required for each parameter) it would not be as reflective of a 'real' program with the inherent difficulties of field monitoring. It should be noted that the impact on the Blackburn Lake analysis by refining the total number of events (using the 'rules' described in Section 3.2) would be to end up with 36 rather than 45 which would reduce the validity of the bootstrapping method (Sandy Creek would reduce from 94 to 32, Cressy St from 104 to 5 (Section 3.6)). The results from this analysis therefore should be seen as an 'upper' range on the number of events required to reduce the uncertainty to approximately 30%.

Log transformation improved the distribution in all cases (Table 3.10).

Table 3-9 Data sets used for uncertainty versus number of event analysis

Catchment Characteristics				**Number of events, N**			**Reference**
Catchment (Location)	**Area (ha)**	**Climate**	**Land use**	**TSS**	**TP**	**TN**	
Blackburn Lake (Melbourne)	221	Temperate	Urban mixed use	45	45	45	(RossRakesh et al. 1999)
Sandy Creek (Brisbane)	227	Sub-tropical	Primarily residential	94	73	73	(City Design 2003)
Cressy Street (Brisbane)	107	Sub-tropical	Primarily residential	104	99	93	(City Design 2003)
Lund (Sweden)	0.027	Cool temperate	Carpark	68			(Deletic 1998)

Table 3-10 Tests of normality for data sets used to determine number of events

Catchment	**Pollutant**	**Shapiro-Wilk 'W'**		**Shapiro-Wilk 'W' (log-transformed data)**	
		Statistic	Significance*	Statistic	Significance
Blackburn Lake (Melbourne)	TSS	0.762	<0.001	0.954	0.074
	TP	0.759	<0.001	0.962	0.144
	TN	0.799	<0.001	0.964	0.176
Sandy Creek (Brisbane)	TSS	0.726	<0.001	0.989	0.657
	TP	0.533	<0.001	0.930	0.001
	TN	0.796	<0.001	0.963	0.032
Cressy Street (Brisbane)	TSS	0.719	<0.001	0.986	0.323
	TP	0.593	<0.001	0.990	0.702
	TN	0.921	<0.001	0.988	0.548
Lund (Sweden)	TSS	0.785	<0.001	0.980	0.328

*Significance of the Shpiro-Wilk statistic in previous column, accepted if >0.05

There is considerable difference in the variability with respect to the number of events sampled between the different sites, with the number of events required to achieve a confidence interval equivalent to 30% of the mean ranging from 50 to 7 for TSS (Figure 3-3), 30 to 50 for TP (Figure 3-4) and 20 to 40 for TN (Figure 3-5). The results for Blackburn Lake were considered most representative in terms of the quality of the data (monitoring was carried out by a specifically dedicated research team, as opposed to contractors and consultants in the Brisbane catchments) and the use of TSS rather than turbidity as in the case of Lund. Being in Melbourne, the data set was also the most representative of the hydrological conditions of the proposed sites. The Blackburn Lake curves showed that approximately 50 events were needed for TSS, 30 for TP, and 20 for TN to achieve a 95% confidence interval equivalent to no more than 30% of the estimated mean. These figures were therefore adopted as the target number of events to be captured for the Melbourne monitoring program.

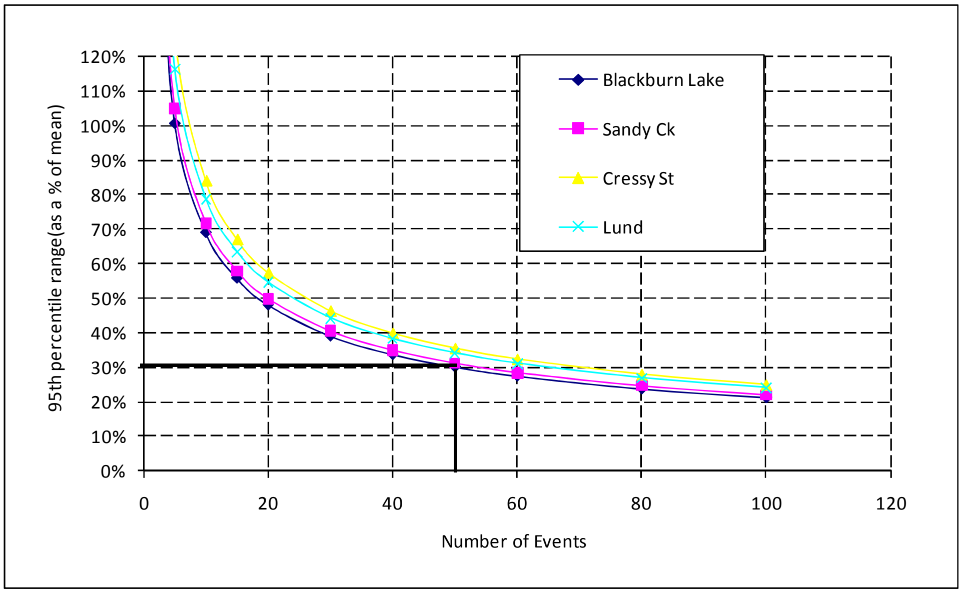


Figure 3-3 **Number of events versus 95th percentile confidence interval for TSS EMCs (heavy black line shows the 30% uncertainty discussed in Section 3.1.1.5)**

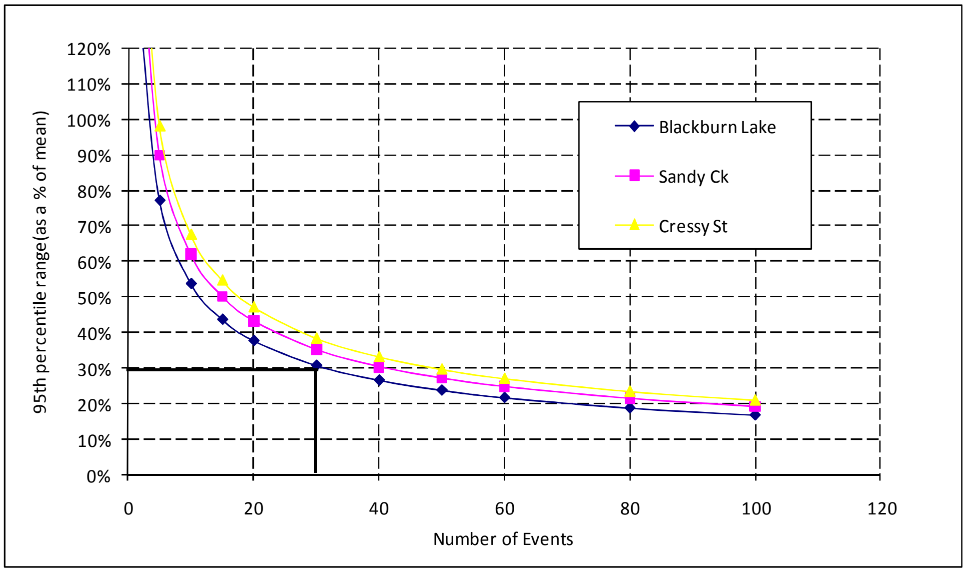


Figure 3-4 Number of events versus 95th percentile confidence interval for TP EMCs (the heavy black line shows the 30% uncertainty discussed in Section 3.1.1.5)

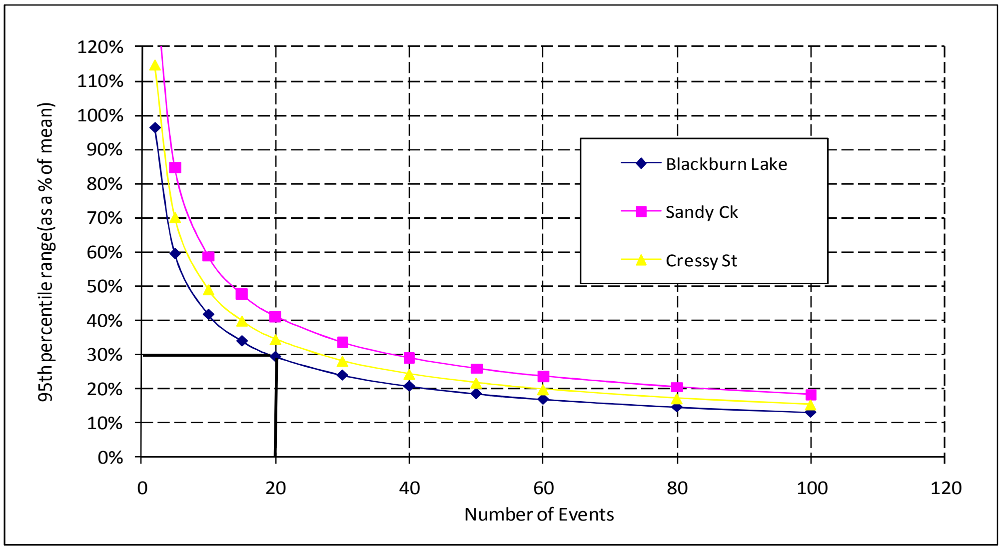


Figure 3-5 Number of events versus 95th percentile confidence interval for TN EMCs (heavy black line shows the 30% uncertainty discussed in Section 3.1.1.5)

3.6 Discussion of secondary data results

Applicability of data sets to hypothesis

The four data sets selected will all be used to test the central hypothesis presented in this chapter. The sites cover a wide range of climatic conditions and have been monitored using methods that allow the investigation of timestep on rainfall intensity (from 10 seconds at Lund and Belgrade to 6 minutes at Blackburn Lake). There are a number of limitations that have been noted as part of the preliminary data investigation, in particular:

- The data sets have all been obtained using very different monitoring methods. Therefore comparisons, and in turn predictions, based on these sites involve considerable uncertainties. One of the major limitations in stormwater quality modelling is the lack of applicability of models calibrated in one catchment for other catchments (Bertrand-Krajewski et al. 1993). To propose a predictive method would require comparison of catchments with similar characteristics

- There is not a wide range of land uses or catchment sizes covered (two catchments less than 300 m^2 and two approximately 200 hectares)

- Two Australian sites are not specifically designed to test a rainfall intensity hypothesis, the rainfall tip size at Sandy Ck (1mm) and the timestep at Blackburn Lake (6-minutes) place limits on exploring the detail of the key hypothesis

- In the two European sites only turbidity as a surrogate for TSS was sampled, and other pollutants of interest were not measured. The use of turbidity, although producing high correlations with TSS in the case of Lund, introduces increased uncertainty when comparing catchments with other monitoring methods

Variability of data
The EMCs for each of the data sets are log-normally distributed. When the 95th percentile ranges of the EMC around the mean are calculated, the Blackburn Lake data set shows the narrowest range. It is not clear whether this would be caused by catchment characteristics, or by the monitoring program undertaken that may have minimised measurement variability (RossRakesh et al. 1999). The Blackburn catchment also had a section of natural waterway which, as discussed, may have contributed to channel processes affecting solids loadings in particular.

The Brisbane City Council data sets were found to have significant gaps in the monitoring data when investigated. Sandy Creek was reduced from 94 to 32 events, and Cressy Street from 104 to less than 5 when the "rules for an event" were applied. Although this data was not taken specifically for the purpose of investigating pollutant generation, this does suggest inadequacies in the monitoring program. Likewise the European data sets were taken using a different monitoring procedure, and they did not include parameters other than TSS. For these reasons the Blackburn Lake data set was considered the most applicable to a Melbourne situation.

Conclusion from secondary data analysis
The data sets investigated provide some valuable information for consideration about the proposed link between rainfall intensity and the generation of urban pollutant loads. In particular, the short timesteps measured at Lund and Belgrade should provide insight into the effect of model timestep on prediction capacity.

However, the difficulty in comparing data, due to varying collection methods, catchments, and climatic conditions, all reduce the power of the combined data sets. Therefore a monitoring program that addresses these issues was proposed for the Melbourne region, and will be outlined in the following section. The monitoring program aims to test pollutant generation across several catchments of varying land use in the same climatic region.

3.7 Primary data sources: Melbourne monitoring program

Given the scarcity of data that fulfilled all requirements of the study, a monitoring program was established specifically to test the hypotheses detailed in this chapter. The monitoring program aimed to take a representative sample of events from a series of catchments and to provide short interval rainfall, flow, and pollutant data for analysis. The program also aimed to test a wide range of catchment sizes and land uses, but also to get some replication of catchments so that comparisons could be undertaken.

Given the hypotheses to be tested, the program was designed to account for variability in both rainfall intensity and runoff rate at as short a timestep as possible. Monitoring equipment was selected and installed with this as a primary aim. The monitoring procedure was designed to reduce the uncertainty of the data collection wherever possible. Uncertainties will be quantified and discussed in Chapter 8.

A process of catchment identification was undertaken to select sites that fulfilled a number of practical criteria:

- Security of sampling equipment
- Proximity and ease of access to allow grab samples to be taken by the research team
- Relatively stable catchments in terms of land-use (a small amount of anticipated development during the sampling period)
- Availability of local drainage plans (in order to determine catchment size)
- Agreement from local government (if sampling apparatus was to be located on a local government system)

Monitoring Sites

Based on the above criteria, seven sites were chosen for the program (Table 3-11). The sites were selected with a variation in size and land use, but with a bias towards medium size urban zones typical of a south-eastern Australian city such as Melbourne. Therefore three of the seven catchments were medium density residential catchments that make up the majority of land use within Melbourne. There is also a low density residential catchment on the urban fringe, a light industrial catchment, and a high density catchment from the inner eastern suburbs of Melbourne. The last site is a coated aluminium roof from a faculty building at Monash University. All samples were taken from underground systems expect for the Monash Roof site.

All sites are in the eastern and south-eastern suburbs of Melbourne and therefore experience relatively consistent climatic conditions. The average rainfall is approximately 700 mm/year mm/year in the region of interest (based on 30 year averages 1961 -2001, http://www.bom.gov.au) with a slight bias towards winter dominated rainfall.

Table 3-11 Summary of Site Descriptions

Site	Primary Land Use	Area (ha)	Total Fraction Impervious	Easting/Northing of sample point MGA94: Grid Coordinates (Universal Transverse Mercator Projection, easting, northing, zone 55)	Melways Reference (Melways Publishing 2002)
Gilby Rd, Mt. Waverley	Commercial	28.2	0.78	336,100 5,803,500	70 G7
Kilgerron Crt, Narre Warren	Rural residential	10.5	0.20	351,200 5,787,500	110 E11
Sheperds Bush, Glen Waverley	Medium density residential	38.0	0.40	340,800 5,805,300	71 J3
Monash roof, Clayton	Coated aluminium roof	0.046	1.00	336,000 5,803,300	70 F10
Eley Rd, Burwood East	Mixed	186.0	0.46	336,600 5,809,200	61 H5
Madden Grove, Richmond	High density residential	89.2	0.74	324,250 5,811,240	44 G12
Ruffeys Lake, Doncaster	Medium density residential	105.7	0.51	335,500 5,817,100	33 F10

Mt Waverley

The catchment is located in Mt Waverley in the south-eastern suburbs of Melbourne. Land use is light industrial (Figure 3-6). The sampling point was from an underground pipe system. Most lots in the catchment comprise warehouses and carparks with very small amounts of landscaped gardens. Approximately 5% of the total area consists of empty lots, not directly connected to the drainage system. Land cover on these lots is grass with small areas of bare soil. The catchment is gently sloping with an average grade of between 1 and 2%.

Figure 3-6 Aerial photo of the Mt. Waverley catchment, with the Melbourne Water Notting Hill rain gauge position shown at the bottom left of the photo. The blue dot shows the location of the rainfall gauge and the yellow star the flow/pollutant monitoring point

Figure 3-7 Typical light industrial land use in the Mt. Waverley catchment

Table 3-12 Data Summary, Mt. Waverley

Data Summary –Mt. Waverley	
Timestep for rainfall gauge	1 minute (on-site) 6 min (Melbourne Water Notting Hill)
Rain gauge measurement volume	0.2 mm
Distance to rainfall gauge (from catchment centroid)	100 m (on-site) 1.1 km (Notting Hill MW gauge ref 586023)
Timestep for flow measurements	1 minute
Parameters sampled	TSS, TP, TN, Al, As, Cd, Cr, Cu, Fe, Mn, Ni, Pb, Zn
Method of pollutant sampling	Discrete sampling
Catchment size (ha)	28.2
Total fraction impervious	0.78
Land use	Industrial (mostly warehouses)

Narre Warren

The Narre Warren catchment is located on the south-eastern fringe of suburban Melbourne (Figure 3-8). The catchment consists of one paved road with an underground stormwater pipe. The sampling pit is shown in **Figure 3-9**.

Average lot size in the catchment is approximately 3000 square metres. Each property contains a primary residence and paved driveway, and most properties also have garden sheds, garages, or other outbuildings. The area is has no wastewater sewer, with each property having a septic tank or local treatment system. The catchment slopes evenly to the monitoring station at an average grade of 4%. On several site inspections, material from septic systems, including clear evidence of toilet cleaning products, was detected in the stormwater system. This was confirmed by inspections and audits of several properties conducted by the Victorian Environment Protection Agency during the monitoring period.

Figure 3-8 Aerial photo of the Narre Warren catchment

Figure 3-9 Typical landscape and monitoring station at the Narre Warren Catchment

Table 3-13 Data Summary, Narre Warren

Data Summary – Narre Warren	
Timestep for rainfall gauge	1 minute
Rain gauge measurement volume	0.2 mm
Distance to rainfall gauge (from catchment centroid)	250 m
Timestep for flow measurements	1 minute
Parameters sampled	TSS, TP, TN
Method of pollutant sampling	Discrete sampling TSS, Composite TP and TN with speciation of N and P for 8 events.
Catchment size (ha)	10.5
Total fraction impervious	0.20
Land use	Low density residential

Glen Waverley

The Glen Waverley catchment is an established medium density residential area in the eastern suburbs of Melbourne.

The catchment is totally residential with typical block sizes in the range of 500 to 700 m^2. The catchment slopes from west to east on an average grade of 4%. The sampling point was situated where the drain levels out at the bottom of the slope (Figure 3-9).

Figure 3-10 Aerial photo of the Glen Waverley catchment

Table 3-14 Data Summary, Glen Waverley

Data Summary – Glen Waverley	
Timestep for rainfall gauge	1 minute
Rain gauge measurement volume	0.2 mm
Distance to rainfall gauge (from catchment centroid)	550m
Timestep for flow measurements	1 minute

Parameters sampled	TSS, TP, TN
Method of pollutant sampling	Discrete sampling for TSS (combination of discrete and composite for TP and TN)
Catchment size (ha)	38.0
Total fraction impervious	0.40
Land use	Medium density residential

Monash Roof, Clayton
The Monash roof catchment is a coated aluminium roof located on the Monash University campus in Clayton (

Figure 3-11). The roof catchment is 0.046 hectares in area. It has a regular slope of approximately 4%. Sampling equipment was positioned inside the university building. The rain gauge was located on the corner of the roof itself.

Figure 3-11 Aerial photo of the Roof catchment

Table 3-15 Data Summary, Monash Roof

Data Summary – Monash Roof	
Timestep for rainfall gauge	1 minute
Rain gauge measurement volume	0.2 mm
Distance to rainfall gauge (from catchment centroid)	15m
Timestep for flow measurements	1 minute
Parameters sampled	TSS, TP, TN
Method of pollutant sampling	Discrete sampling for TSS (Composite sampling for TP and TN)
Catchment size (ha)	0.046 ha
Total fraction impervious	1.0
Land use	Colourbond roof

Richmond

The Richmond catchment is located in the inner eastern suburbs of Melbourne, and comprises mainly high-density residential housing, with small (< 5%) areas of commercial and light industrial use Figure 3-12 and 3-13). Average lot sizes are 300 to 500 square metres. Isolated infill construction occurred within the catchment during the monitoring period, but was not atypical for the area and hence was not considered unrepresentative for the monitoring program. Small urban parks comprise less than 2% of the catchment. Most of the catchment is flat with slopes of less than 0.1%, although there is a hill along the western edge with slopes of up to 7%. Street sweeping (aimed at gross pollutants) is carried out fortnightly on residential streets in the area and more regularly on the commercial areas.

Figure 3-12 Aerial photo of the Richmond catchment, with the MW Burnley rain gauge position shown at the bottom

Figure 3-13 Pipe outlet at Richmond. The sampling point is 250m upstream of the pipe outlet

Table 3-16 Data summary, Richmond

Data Summary – Richmond	
Timestep for rainfall gauge	1 minute (Bunting Street) 6 min (Melbourne Water, Burnley)
Rain gauge measurement volume	0.2 mm
Distance to rainfall gauge (from catchment centroid)	600 m (Bunting Street) 1.1 km (Burnley MW gauge ref 229621)
Timestep for flow measurements	1 minute
Parameters sampled	TSS, TP, TN, Al, As, Cd, Cr, Cu, Fe, Mn, Ni, Pb, Zn
Method of pollutant sampling	Discrete sampling
Catchment size (ha)	89.2
Total fraction impervious	0.74
Land use	High density residential

Burwood East

The Burwood East catchment is also an established medium density residential area in the eastern suburbs of Melbourne. The sampling point is at the entrance to a Melbourne Water owned retarding basin (**Figure 3-**).

Typical block size is in the range of 500 to 700 m^2. There is also a large (approximately 6 hectares) commercial area and several small playing fields and parks (Figures 3-14 and 3-15).

Figure 3-14 Aerial photo of the Burwood East Catchment with catchment boundary. The flow and pollutant sampling point is shown by the yellow star. The blue dot is the alternate monitoring point and the one-minute rainfall monitoring point is on the right hand boundary of the photo

Figure 3-15 Typical streetscape in the Burwood East catchment

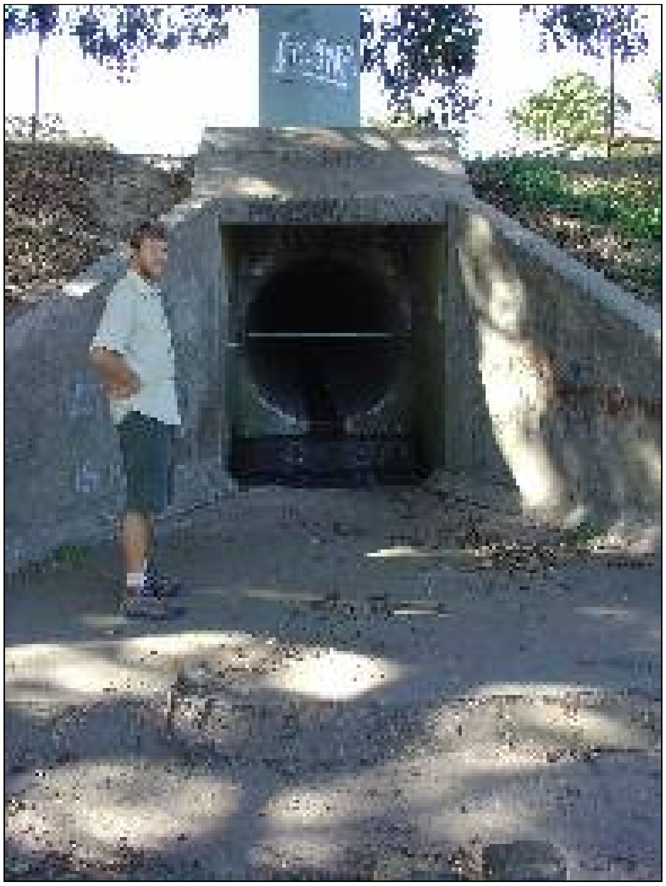


Figure 3-16 Outlet at the Burwood East catchment

Table 3-17 Data Summary, Burwood East

Data Summary – Burwood East	
Timestep for rainfall gauge	1 minute (Peter James Institute) 6 minutes (Eley Rd RB MW gauge ref 229638)
Rain gauge measurement volume	0.2 mm
Distance to rainfall gauge (from catchment centroid)	1 km (Peter James Institute) 1.3 km (Melbourne Water Eley Rd RB)
Timestep for flow measurements	1 minute
Parameters sampled	TSS
Method of pollutant sampling	Discrete sampling
Catchment size (ha)	186
Total fraction impervious	0.55
Land use	Medium density residential

Doncaster

The Doncaster catchment is an established medium density residential area in the eastern suburbs of Melbourne. An indoor shopping centre, Doncaster Shopping Town, surrounded by carparks, occupies about 13% of the catchment at its upstream extremity (**Figure 3-**). The average slope is 5%, and typical lot sizes are in the range of 600 to 950 square metres (Figure 3-18). There are several sporting fields and an area of parkland with native vegetation in the catchment.

Figure 3-17 Aerial photo of Doncaster catchment, with the Eastern Golf Club (and alternate rain gauge shown by the blue dot) visible on the left of the photo. The yellow dot in the top right shows the flow, pollutant and rainfall gauge

Figure 3-18 Typical streetscape from the Doncaster catchment, the multi-storey building in the top right of the photo is Doncaster shopping town at the upstream extremity of the catchment

The sampling site is immediately upstream of the pipe outlet (**Figure 3-**). There is a small “basket” type litter trap upstream of the sampling point that is not regularly maintained and therefore is expected to have little influence on pollutant loads. Because it is not continuously submerged it is also unlikely to have significant influence on pollutant speciation.

Figure 3-19 Pipe outlet at Doncaster (sampling point is 15 m upstream)

The Doncaster catchment itself has already been studied in terms of the composition of nitrogen in its stormwater, during both dry and wet weather (Taylor et al. 2005). The first 14 events from the data set formed the inflow data in that research. Of those 14 events, only six have six-minute rainfall data from the Melbourne Water gauge at the nearby Eastern Golf Club. As can be seen in Figure 3-19 there is some backwater from the wetland itself that may have influenced the monitoring results in some cases, particularly in large events that filled the wetland. When the current study started the monitoring point was moved further upstream in the pipe however in large events the backwater still impacted on the flow measurement. This is likely to add uncertainty to the flow measurement at the site. This effect has been quantified in a separate study at the site (McCarthy 2008) which demonstrated that the flow volume uncertainty at the site was 38%, greater than that at other sites in the study.

Table 3-18 Data summary – Doncaster

Data Summary – Doncaster	
Timestep for rainfall gauge	1 minute (on-site) 6 minutes (Eastern Golf Club MW gauge ref 586010)
Rain gauge measurement volume	0.2 mm
Distance to rainfall gauge (from catchment centroid)	700m (on-site gauge) 1.7 km (Eastern Golf Club)
Timestep for flow measurements	1 minute
Parameters sampled	TSS
Method of pollutant sampling	Discrete sampling
Catchment size (ha)	105.7
Total fraction impervious	0.51
Land use	Medium density residential

3.8 Monitoring Procedures

Rainfall measurement

Rainfall was collected using 0.2 mm tipping gauges at a timestep of one minute. Rainfall tips were assigned to the minute in which it was recorded. The gauges were located as close as possible to the centroid of the catchment at a site that was protected as much as possible from interference (within a private property or out of site from roads and footpaths). The distance from the centroid of the catchment to the rainfall gauge is given in the site summary for each site.

In all cases except Narre Warren and Glen Waverley there is also a Melbourne Water rain gauge within 2 km of the catchment centroid with a relatively continuous set of rainfall data. This gauge network records 6-minute rainfall data that are stored on a data base at Melbourne Water. The gauge numbers associated with the Melbourne Water gauges are also given in the site summaries.

Flow measurement

Sigma 900 autosamplers and Sigma 950 flowmeters (Figure 3-0) were installed in locked cabinets at each of the pipe outlets or at a convenient drainage pit. The runoff rate was recorded at one-minute timesteps using the cross-sectional area (calculated using the ultrasonic pressure transducer) and the velocity (using the Doppler ultrasonic sensor). The flow sensor was positioned in the pipe to minimise both backwater and sediment/litter buildup effects. The sensor was positioned 2 to 4 cm above the pipe invert to prevent clogging from bed loads of sediment. As can be seen in Figure 3-13, the sensors and associated cables are prone to collecting leaf litter and other debris, therefore inspection and cleaning of the sensors was undertaken after each wet weather event.

Figure 3-20 Autosampler setup

Figure 3-13 Sampling point with associated leaf litter

Because of the small flows associated with the Monash roof, the flow was calculated using a V-notch weir connected to a downpipe (Figure 3-14).

During the monitoring program the rainfall results were continuously evaluated against the runoff results to assess functioning of the experimental system. This led to the following adjustments during the monitoring period:

- Because unreliable flow measurements were recorded at Burwood East on several occasions (discovered through small recorded flow when compared to large rainfall totals), the autosamplers and flowmeter were changed on several occasions in an attempt to identify the problem. However, the issue was ongoing throughout the program, resulting in only a small number of events being captured at the site

- The rainfall runoff relationships recorded at the Monash roof often were unrealistic (more runoff than rain) during the first few events. The rainfall gauge was initially several hundred metres from the sampling site on top of another university building. Whilst this might seem like a small distance, the catchment is very small (459 m^2), and thus spatial variability of rainfall becomes critical. To rectify this, the rainfall gauge was moved to the roof itself.

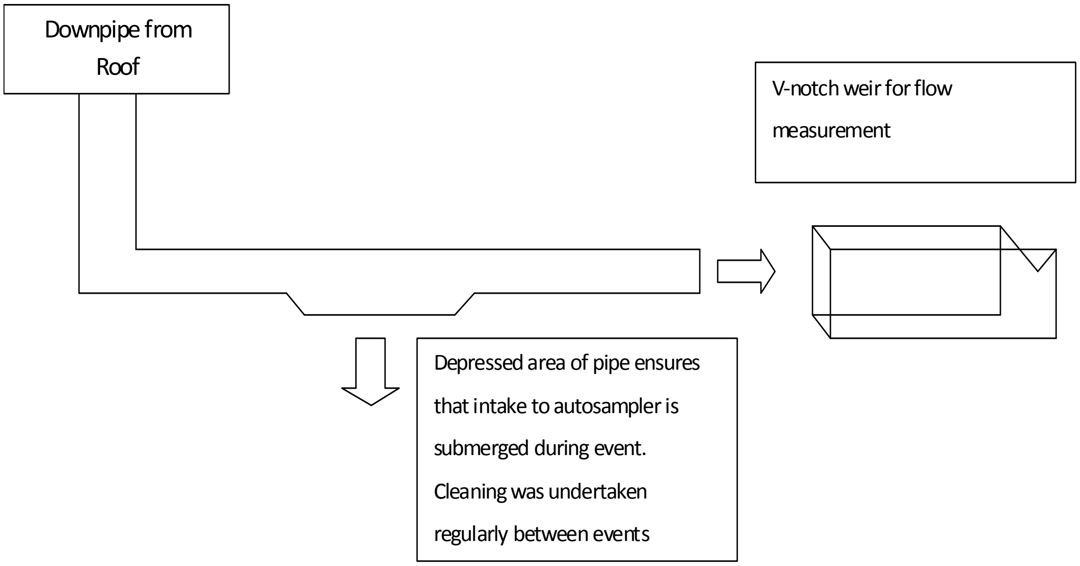


Figure 3-14 Monitoring setup at the Monash roof site

Pollutant sampling

Dry weather runoff pollutant samples were taken after a period of three days without rain at the pipe invert. All sites were sampled during dry weather (except for the roof site) by simply holding a sample bottle in the flow at the bottom of the pipe or at the pipe outlet.

For wet weather events, flow sensors (as shown in Figure 3-) were connected to the autosamplers, and programmed to trigger the samplers after a set increase in flow (which varied depending on the catchment). Samples taken were controlled by a pulse sent from the flowmeter when a predetermined volume had passed.

Initially the autosamplers were programmed to achieve maximum resolution for a typical 1 in 3 month average return interval (ARI) event. To determine the sampling frequency a three-month ARI storm volume was estimated using MUSIC (CRCCH 2003). Of the 24 bottles available, ten were used uniformly in the first 30% of this flow volume, ten for the next 40%, and four bottles for the last 30% of the event. This enabled a balance between definition for the rising limb of the event and capturing larger events. This method was refined throughout the sampling period by adjusting the expected volumes based on experience with the catchment.

The method varied slightly for the Monash roof site where, because of the catchment size, only small flowrates were recorded. In that case a small depression was made in the bottom of a horizontal pipe that carried the flow (see Figure 3-14). The sampler was positioned in the depression so that it would always be inundated during an event. The depression was cleaned out regularly to avoid deposition of solids.

Where possible, discrete sampling was undertaken for all pollutants. Due to cost issues, however, some of the TP and TN samples were gathered using a composite approach where a known volume of water was collected at programmed intervals during the event. Where discrete samples were taken the pollutant concentrations was linearly interpolated between samples to determine a concentration at each recorded timestep. Pollutant loads were calculated by multiplying the flow by the concentration at each timestep in the hydrograph. The pollutant load for an event was then calculated using the “rules” defining an event to determine the starting and ending points (see Table

3-1). When composite sampling was undertaken the event mean concentration (EMC) was multiplied by the volume to determine the load.

Samples were analysed at the NATA-accredited Water Studies Centre laboratory at Monash University. Standard Methods were used for TSS, TP and TN (Greenberg et al., 1999).

Calculation of catchment fraction impervious

The GIS system MapInfo (MapInfo Corporation 2002) was used in conjunction with underground drainage plans and one metre contours to calculate the areas for each of the catchments. The land use within each of the catchments was then defined using the planning scheme zonings (Department of Sustainability and Environment website, www.dse.vic.gov.au – planning schemes online 2005).

An indicative fraction impervious was selected for each land use within the catchment based on examination of aerial orthophotos and site inspections. The weighted average (based on area) of these land uses was then used to calculate a total fraction impervious for the catchment. One metre contours were also used to calculate the average slope for each catchment.

3.9 Description of the primary data set

Table 3-19 summarises the data that have been reliably collected in this monitoring programme. The data set is very comprehensive. Approximately 250 events were captured across six catchments, all with short timestep rainfall and flow data. TSS, TP and TN as well as nutrient speciation were captured at all sites. Corresponding low flow data was captured for all sites with base flow.

The data set provides a short-interval rainfall and flow data across a range of catchment sizes and urban land uses. The data will provide a level of consistency and confidence in the testing of the hypotheses that was not possible with the secondary data sets. It will also allow the comparison of these sites with the secondary data sources to examine if data quality is a factor in previous studies aimed at establishing a common driver for pollution loads.

Table 3-19 Summary of events collected (Note: the table includes all events for which flow and pollutant data was considered reliable, some of these events do not have corresponding rainfall data and therefore the numbers here may be slightly different to those presented in the further analysis)

Catchment	Pollutant	Number of events collected	Dry weather samples collected (Melbourne 2003/05 monitoring program)	Comments
Mt. Waverley	TSS TP TN	49 47 47	14 14 14	
Glen Waverley	TSS TP TN	19 17 17	16 16 16	Problematic site, first with lightning strike on the rain gauge then with build up of sediment at the monitoring site.
Narre Warren	TSS TP TN	41 18 18	16 16 16	Speciation of nitrogen for 8 samples. Rainfall gauge failed during several events.
Monash Roof, Clayton	TSS TP TN	30 30 30		
Burwood East	TSS TP TN	23* 17* 17*	16 16 16	
Richmond	TSS TP TN	40 39 39	10 10 10	
Doncaster	TSS TP TN	54	10 10 10	14 original events conducted as part of study by Taylor(2005)

* Only six-minute rainfall available for 13 of these events at Burwood East

Quality of the primary data

The problems encountered during the monitoring can be grouped into five main categories:

- Equipment failure
- Catchment behaviour
- Adaptive management of the program
- Errors in measurements of flows and pollution concentrations
- Achieving targeted number of monitored events

On several occasions equipment failed or malfunctioned, causing storm events to be missed, or suspect data to be collected. Some of these, such as a lightning strike on the rain gauge at Glen Waverley, were unavoidable. Building redundancy into the program – such as having a backup rain gauge – would have meant that flow data from that period could have been used despite the malfunction. Unfortunately, due to financial constraints, that was not the case at Glen Waverley. Other situations however such as the build up of debris on flow transducers could have been better managed by a formal inspection process during the monitoring.

Assessment of the catchment behaviour, especially early in the program was also important to ensure quality data. As discussed, Burwood East catchment experienced several non-rainfall related flow periods, often late at night, which suggested that inappropriate activities were occurring in the catchment. Narre Warren also experienced significant periods of highly polluted flow from septic systems which would skew results if not identified. The "rules for the event" defining a runoff co-efficient within an expected range was the most important tool to ensure quality. Less than 5% of events captured were discarded during the program (although slightly more at the Roof site due to the inappropriate positioning of the rain gauge initially). Equipment malfunction was the most common reason was the most common reason for discarding events.

Issues such as the repositioning of the rain gauge for the Roof site (see Section 3.7) were positive examples of how the program could change based on the initial results. Other examples such as the backup raingauges at Burwood East and Doncaster also were used adaptively during the program.

Errors in measurements of pollution concentrations

Uncertainties in flow and stormwater pollutant data have been widely reported in the literature (Arnberg-Nielson and Harremoes 1996; Ahyerre et al. 1998; Bertrand-Krajewski et al. 2002). Masimovic (1986), cited in Ahyerre et. al. (1998), report flow uncertainties of 5 to 25% depending on the method used. For a pressure transducer system (as used in the current study), flow uncertainties of approximately 10% have been published (GeoSyntec Consultants and ASCE 2002). Combined flow and pollutant sampling uncertainties are usually in the order of 30% for TSS (Bertrand-Krajewski et al. 2002).

McCarthy et. al (2008) used four of the sites from the primary data set to investigate the uncertainties involved in flow measurement (as part of a larger study which looked at E. coli generation in stormwater). These sites were Mt. Waverley, Doncaster, Narre Warren and Richmond. The study found flow volume uncertainties of 10%, 38%,13% and 15% respectively.

Most published uncertainty data focuses on TSS. In light of the trend for decreasing TSS concentrations worldwide that have been reported (see Chapter 2) and the focus on sampling procedures as a cause for this, several procedures were undertaken to confirm the total experimental uncertainty associated with sampling procedures:

- Experiments were conducted on the sampling variability caused by the height of the monitoring probe within the drainage system This is directly aimed at addressing the issue of "bed load" sediment being sampled and recorded erroneously as TSS. The procedure was undertaken for TP and TN in addition to TSS. Samples were taken from two points, one 3 to 5 cm above the pipe invert (typical for the pollutant sampling in the programs being considered) and one from approximately 13 to 15 cm above the invert. Five repetitions were taken for TSS, TP, and TN—all at one site

- The variation in pollutant concentration analysis from different laboratories was tested by sending identical samples to three different laboratories

- To determine the uncertainty in pollutant analysis, simultaneous grab samples were taken during wet weather flow and were sent to three independent laboratories. This process was repeated for six samples

Pollutant samples, as has been reported previously, were higher at the bottom of the pipe in all cases (except for TN in one case). Differences between the low and high are similar those reported in Ahyerre et al. (1998), and less for TN than TSS. TP had the highest average difference although the scatter in TP results was high (although the resolution of the TP results was low with only one significant figure, therefore the apparent differences could just be a function of the analysis). Average differences for the sampling are 13, 16, and 3 percent for TSS, TP, and TN respectively (Table 3-20) which is within the range of previous studies. As expected, nitrogen, which has the highest dissolved percentage of composition of the parameters tested (Taylor et al. 2005) is more consistent through the water column.

Table 3-20 Results of depth profile experiment

Pollutant (mg/l)	Sample number	Top	Bottom	% difference
TSS	1	28	30	7
	2	28	33	18
	3	26	28	8
	4	27	31	15
	5	27	32	19
TP	1	0.09	0.10	11
	2	0.06	0.08	33
	3	0.08	0.08	0
	4	0.06	0.08	33
	5	0.08	0.08	0
TN	1	0.70	0.82	17
	2	0.62	0.62	0
	3	0.62	0.60	3
	4	0.60	0.62	3
	5	0.62	0.62	0

Table 3-21 Results of sample analysis by different laboratories

			Laboratory (individual samples)		
Pollutant (mg/l)	**Site**	**Range**	**1**	**2**	**3**
TSS	1	3	43	45	42
	2	6	36	42	38
	3	9	30	35	24
	4	6	29	35	32
	5	4	18	19	14
	6	16	25	32	16
TP	1	0.20	0.06	0.04	0.24
	2	0.26	0.07	0.02	0.28
	3	0.61	0.13	0.12	0.73
	4	0.15	0.14	0.13	0.28
	5	0.27	0.18	0.14	0.41
	6	2.39	0.26	0.18	2.57
TN	1	0.4	1.5	1.7	1.3
	2	0.3	1.5	1.4	1.2
	3	0.7	2.4	2.4	1.7
	4	0.7	2.5	2.3	1.8
	5	0.1	1.3	1.2	1.2
	6	0.5	1.7	2.2	2.1

When the laboratory testing is considered the difference are generally within the ranges quoted by Ahyerre et al. (1998) with ranges of up to 40% for TSS. Laboratories 1 and 2 are also generally in agreement with TP and TN, although the differences are larger than for TSS. Laboratory 3 is significantly different for TN and TP, with TP samples in particular varying wildly. Although not conclusive, and of some concern, these results are within the range of those used for previous analysis.

For the purposes of program design, these experiments confirm the previously published work that proposes a total uncertainty of approximately 30% for pollutant load estimation from stormwater events. Therefore a guiding figure of 30% monitoring uncertainty is proposed as a target figure for the analysis of variability in design of the Melbourne monitoring program (Section 3.6)

Achieving targeted number of monitored events

The original goal of the monitoring program was to sample 50 events for TSS at most sites, and this was achieved only at Doncaster, although Mt. Waverley, Narre Warren, and Richmond all reached over 40 events. Likewise, most sites totalled about 20 events for TP and TN where the original goal was at least 20 to 30.

In particular the difficulties with the Eley and Glen Waverley catchments have reduced the strength of the data set. The program was originally designed to obtain some repetition of results, particularly for the medium density residential land which makes up the greatest land use by area in Melbourne.

In both the examinations of monitoring uncertainty, the sample analysis and the flow profile, the results are equal or greater than other relative uncertainty figures reported in the literature. In particular the laboratory analysis of TN and TP shows great differences, particularly in the case of laboratory number 3. Given these results and the shortfall in numbers of events captured in some cases care will need to be taken when comparing predicted results with the monitoring data.

4 PRELIMINARY DATA ANALYSIS

4.1 Introduction

Chapter 4 provides an overview of the data sets collected both through the literature (termed the “Secondary data sets”) and through the purpose-designed monitoring programme (termed the “Primary data sets”). Based on the studies described in Chapter 3 the data will be examined and compared with previous descriptions of stormwater behaviour.

The chapter will present a statistical analysis of the rainfall, flow and pollutant data. Three major aims of this chapter are:

- To analyse how representative the data collected is of typical conditions at each catchment.
- To present the data set in a form suitable for comparison with other published data.
- To identify any abnormalities in the data collected that may influence the testing of the hypothesis presented in chapter 2.

As described in Chapter 3 and by many authors (Deletic et al. 1997; Ahyerre et al. 1998) the methods and scope used in the monitoring program are important underpin conclusions drawn. Factors such as the spatial and temporal resolution in particular impact on the confidence of the data outputs.

In the current study stormwater quality for each site, described using Event Mean Concentrations (EMCs) and pollutant loads, is compared to various predictors of stormwater quality that have commonly been used in the literature (e.g. rainfall volume and intensity, runoff volume and rate, etc). The results of this analysis are presented and compared to past studies.

Lastly the chapter will introduce the concept proposed in Chapter 2, regarding potential relationships between event loads and the rainfall/runoff functions. Chapter 5 will further develop the central hypothesis and test the strengths and limitations of the load based relationships.

4.2 Background

Given sufficient monitoring data it is theoretically possible to accurately represent the long-term runoff loads from a catchment using statistical methods. For example, pollutant export from urban catchments is often described in terms of the statistics such as the event mean concentration (EMC) - the load of pollutant over an event divided by the volume of flow during the event (Sharpin 1995; Smullen et al. 1999; Brezonik and Stadelmann 2002). When extensive records of urban stormwater EMCs have been analysed they have usually been shown to be log-normally distributed (Driscoll 1986; Duncan 1999; Fuchs et al. 2004). Event Mean Concentrations have been shown to vary significantly between catchments, or for different events within the same catchment (Chiew and McMahon 1998), making prediction of water quality from a given catchment more difficult.

Despite this, there have been many predictive methods and models such as MUSIC (CRCCH 2003), which uses concentrations stochastically generated from a log-normal distribution, to predict long-term pollutant loads, using either estimation of EMCs, or stochastic generation of concentrations based on EMC and its standard deviation. Unfortunately, local data to support the use of these models is rarely available, and even when available, the data is often of questionable quality (Chiew et al. 1997; City Design 2003; Francey et al. 2004).

Whilst these models can be used to predict long term loads, subject to the constraints on local data, they do not provide any insight into the behaviour of pollutants over short timesteps, or inform about the processes responsible for pollutant generation and /or washoff.

Given that the variation in urban runoff quality is considerable, many researchers have taken a more deterministic approach and attempted to explain the variance in concentrations recorded during storm events through relationships with rainfall, flow volumes or antecedent conditions (Desbordes and Servat 1984; Driver 1990; Duncan 1995). As discussed in Chapter 2, many models are available which use various physical parameters to predict pollutant behaviour. The primary drivers behind pollutant mobilisation, however, are contentious (Bertrand-Krajewski et al. 1993; Chiew et al. 1997). Factors such as rainfall (intensity and volume), runoff (intensity and volume), antecedent dry days and combinations of these have been used, but a explanation that works across many catchments, with varying climates and land uses, has been elusive (Duncan, 1999; Kanso and Tassin, 2004).

Several studies which describe pollutant export (GHD and EPA Vic 1981; City Design 2003) or attempt to explain pollutant behaviour (Srinanthakumar and Codner 1993; Chiew and McMahon 1998; Walden 1999; Vaze and Chiew 2003) have been undertaken using Eastern Australian data sets. Results from these studies will be compared with the current study, in order to identify significant similarities, differences, or emerging trends.

4.3 Method

Data was considered in four sections:

- Rainfall results for each catchment
- Runoff results (including rainfall runoff relationships)
- Influence of hydrological indicators on Pollutant concentrations
- Pollutant loads (including relationships with rainfall and runoff functions)

Initially the rainfall, flow and pollutant data sets were assessed for their consistency and representativeness. Maximum and average flow rates and intensities were used in order to compare hydrological conditions between sites and to assess the range of events captured at a particular site. This was done in order to identify any catchment specific behaviour that was unexpected, and that would potentially require further explanation, testing or analysis.

Rainfall

Two aspects of the rainfall data were considered; the catchment-based statistics in order for cross-catchment comparisons, and the comparison of selected sites with long term records. For each event the total event rainfall, the average rainfall intensity over the event and the maximum rainfall intensity in a 6-minute timestep were calculated. The 6-minute timestep was chosen to reflect the smallest timestep that could be used for all catchments (Blackburn Lake was only available in 6-minute timesteps). The range of maximum rainfall intensities was also presented for comparison between catchments.

These statistics were selected to reflect the range of hydrological conditions that occurred during the monitoring programs. They are also indicators that have been used by a wide range of previous studies (Driver and Troutman 1989; Walden 1999) to attempt to predict pollutant loads and concentrations, and thus provide a useful basis for comparison.

For the Melbourne events where long term data was available, the approximate average recurrence interval (ARI) of each event was calculated using methods described in Australian Rainfall and Runoff (1997). Each event recorded in the current study was compared against a historical intensity-frequency-duration curve for that site at 1, 5, 20 and 100 years recurrence intervals. Rainfall data for two sites was available for this analysis, Mt. Waverley and Richmond, due to their proximity to a Melbourne Water rainfall gauge with a comprehensive set of six-minute rainfall data. Reliable rainfall data was available for approximately 80 years for both of these sites. This procedure was carried out using the Melbourne Water hydrological database (HYDSYS). This allowed a comparison between the events captured during the monitoring program and the long-term record for the area.

Runoff

Flow was measured continuously at the same point as pollutant samples were taken. This is presented in terms of the event peak, expressed in l/s for an absolute comparison between sites, and expressed in runoff depth (mm/hr) to allow an area-scaled comparison between sites.

The average value of all event maximum flow rates and runoff depths, and the range of these maximum values, was used to compare the flow between sites.

The relationships between total event rainfall and total event runoff volume were compared and graphed in the form of rainfall/runoff plots. Rainfall/runoff plots are widely used to assess the quality of runoff monitoring as well as to investigate the impact of pervious/impervious area runoff (Boyd et al. 1994; Shuster et al. 2005). A poor or inconsistent relationship between rainfall and runoff would indicate probable monitoring problems and would have implications for the data quality as discussed in Chapter 3. There are several possible explanations for an inconsistent relationship between rainfall and runoff:

- The catchment may respond differently to different rainfall patterns, for instance a catchment with significant pervious areas may experience different levels of surface runoff depending on the intensity and/or duration of the rainfall. It is expected that urban catchments with large impervious fractions would be largely buffered from this effect (Boyd et al. 1994)

- The rainfall may have been inconsistent over the catchment (due to spatial variability). This mostly affects large catchments, but has also been demonstrated for small urban catchments (Jakeman and Hornberger 1993; Ogden and Julien 1993)

- Antecedent moisture conditions may alter between events

- There may be a calibration or other problem with the rainfall or flow recording equipment

Despite these potential factors, a very high correlation was expected for rainfall runoff plots for small urban catchments, given the relatively well designed and managed monitoring setup and a representative spread of events.

Pollutants

Distributions of Wet Weather EMCs

Before statistical interpretation of the data was considered valid, an examination of the distribution of the data was necessary. EMC data has been shown to be log-normally distributed (Duncan 1999). The concentration data was therefore tested for normality using untransformed and log transformed data using the Shapiro-Wilk distribution test (see Chapter 3).

Wet Weather Events

Examination of the concentration data was undertaken using Site Mean Concentration (SMC) and Event Mean Concentration (EMC). Initially pollutant loads were calculated by linearly interpolating the measured pollutant concentrations across each timestep of the event.

The SMC of each pollutant was calculated by dividing the total mass of pollutant across all events by the total runoff volume. This provides a parameter for comparison between sites. If discrete sampling was available for an event the EMC for each event was calculated by dividing the pollutant load by the event volume. If composite sampling was undertaken that value was adopted as the EMC. Where available, the speciation of nitrogen and the filterable reactive phosphorus was also recorded.

Both EMC and SMC methods are widely used in the literature and therefore were both reported. The SMC parameter has an advantage over the EMC, particularly when describing long term behaviour, because large events, which may carry a disproportionate amount of the total load, are weighted more heavily than small events. The EMC however allows comparisons between individual events, and their variability to be assessed.

The exception to this method was the published Lund and Belgrade data sets (Deletic 1998) where turbidity was used as a surrogate for TSS. In these studies a relationship between turbidity had been derived for each site (see catchment description in Chapter 3) based on site specific calibrations. Based on these calibrations TSS loads were calculated and reported for each event. For these sites the event loads reported in the literature (Deletic 1997) were used.

Metals and some ionic components were also measured at three of the sites as described in Chapter 3; Mt. Waverley, Richmond and the Monash Roof. Because the number of sites and events was limited the statistical strength of the results was not as great as that for other pollutants. The EMCs therefore were not compared with hydrological variables but were used to compare EMC values between catchments.

Pollutant concentrations: dry weather/baseflow

Dry weather (base flow) sampling was also undertaken as part of the study (samples were taken periodically after periods of 3 days without rainfall). Basic statistics of the dry weather samples for TSS, TP and TN are reported.

Although dry weather concentrations are unlikely to be related to the hydrological models being considered here, they are an important descriptor of catchment conditions and help fulfill the secondary aims of this study, the development of an extensive, local stormwater quality data set. SMC from the wet weather events was then plotted against the Dry Weather mean values and compared to previous studies.

Correlations between EMCs and Wet Weather Event Characteristics

Correlation analysis between the log-transformed EMC data and a variety of storm event descriptors was undertaken:

- Total event rainfall – By using the “rules for an event” defined in chapter 3 the rainfall was summed (in mm) over the time period corresponding to the EMC calculation

- Total event volume - By using the “rules for an event” defined in chapter 3 the runoff volume was summed (in litres) over the time period corresponding to the EMC calculation

- Maximum rainfall intensity – The maximum rainfall recorded over a 6 minute timestep for all sites (all sites were “lumped” up to 6 minutes as it was the minimum timestep recorded at some catchments) was used to identify the maximum intensity, and then transformed into mm/hr

- Average rainfall intensity – the total event rainfall divided by the event duration

- Maximum flow rate – The largest flow recorded over a one minute timestep

- Antecedent conditions (Number of preceding days with less than one mm rain)

Maximum rainfall intensity, average rainfall intensity and flow rate have often been used to represent the energy input from the storm event (Desbordes and Servat 1984; Driver and Troutman 1989; Walden 1999). A higher EMC may be expected in a more intense storm if maximum values are a good descriptor of the process. Likewise the antecedent dry days may be an important descriptor if the buildup of pollutants on the surface during dry periods was a major factor in pollutant generation. While these factors have been extensively used, and successfully in some isolated cases (Driver and Troutman 1989), they have not been found to be applicable on a widespread basis.

Data from each site was examined for cross correlations between pollutants. In particular the correlation between phosphorus and TSS concentrations which has been reported (Sartor, Boyd et al. 1974; Xanthopoulos and Hahn 1993) or assumed in published studies was tested.

For all correlations the Spearman non-parametric correlation coefficient (r_s) (Fowler et al. 1998) was used in addition to the standard Pearson coefficient of determination. This was used to account for the very large data points that may dominate the coefficient of determination.

The formula for r_s is:

$$r_s = 1 - \left[\frac{6\Sigma d^2}{\left(n^3 - n\right)} \right]$$ **Equation 4-1**

where n is the number of units in a sample and d is the difference between the ranks of the individual samples.

Relationship between event load and explanatory variables

The method used to calculate loads has been described in Chapter 3. The calculated pollutant loads were compared with a range of potential explanatory variables, including total runoff, total rainfall, and a *rainfall function* and *runoff function*. The central hypothesis of this study (Chapter 3), the approach of Duncan (1995), tested by Vaze and Chiew (2003), which used a power of the intensity summed over short time-steps, is proposed to represent the total energy supplied through rainfall. The equation given previously may be expressed in the following form to estimate both the total load of a pollutant over a storm event and the pollutant flux within an event at a given site:

$$Load = WTA\sum_{i=1}^{n} I^b$$ **Equation 4-2**

where *Load* is the load of pollutant produced over the period of interest (units of mass), I is rainfall intensity (units of depth/time) measured at short time periods (6 minutes or less),T is the timestep of the monitoring (units of time), A is the catchment area (units of length^2), n is the number of equal time intervals for which I was measured over the period of interest, W is a calibration coefficient dependent on the pollutant and catchment characteristics, and b is a calibration coefficient dependent only on the pollutant. The intensity function is referred to as the rainfall function.

If the runoff rate (R) is substituted for the intensity at each timestep the equation becomes

$$Load = CTA\sum_{i=1}^{n} R^d$$ **Equation 4-3**

where C is a calibration coefficient dependent on the pollutant and catchment characteristics, and d is a calibration coefficient dependent only on the pollutant. The runoff rate term is referred to as the runoff function.

Initial screening was carried out through regression analysis of the event load of each contaminant against each explanatory variable, for each of the sites, and tabulating the coefficient of determination (R^2) for each combination. For the rainfall function and the runoff function, the power coefficients *b* and *d* were optimised iteratively to achieve the highest possible R^2 at each site.

Total event rainfall is therefore represented when the b term in Equation 4 -2 is one. Likewise, total event runoff volume is represented when the d term in equation 4.3 is one. The first stage of this analysis is to compare regression results from the rainfall function with those of the runoff function, the total event volume and the total event rainfall.

The behaviour of the optimised values of *b* and *d* will be discussed in Chapter 5. The coefficient of determination places the most weight on the larger values in the dataset (Legates and McCabe 1999). Because event loads are being considered this was seen as an appropriate approach; other methods of testing "goodness of fit" are used in subsequent chapters.

4.4 Results

Rainfall

A large range of rainfall totals and intensities were captured during the monitoring program (Table 4.1). The mean and median rainfall intensity data are broadly similar for the Melbourne and Belgrade catchments. The high mean value for the Brisbane catchment of Sandy Creek reflects the high intensity sub-tropical rainfall; conversely the low mean value for Lund reflects a temperate pattern with many days of low intensity drizzle.

Mt. Waverley and Richmond both had high values of maximum rainfall intensities (76 and 68 mm/hr respectively). These events both represent "thunderstorms" late in 2004 that were above the 1 in 20 and 1 in 100 year ARI for their respective areas (Figures 4.1 and 4.2).

Table 4-1 Summarised Rainfall data for all catchments

Site	Number of events	Range of event total rainfall (mm)	Mean maximum rainfall intensity (mm/hr)***	Median maximum rainfall intensity (mm/hr)***	Range of event maximum rainfall intensities (mm/hr)***
Primary Data Sets					
Mt. Waverley	49	1.2 – 38.6	12.3	10	2 -76
Glen Waverley	32	0.6 – 21.0	9.25	6	2 – 26
Narre Warren	38**	0.4 - 110	12	10	3 - 34
Monash Roof	26	1.2 - 15.8	10.8	10	2 -28
Burwood East	23	2.0 – 32.6	13.1	8	2 - 50
Richmond	40	1.2 – 40.8	12.9	10	4 -68
Doncaster	54	1.6 – 95.8	10.1	7.0	4 -34
Secondary Data Sets					
Blackburn Lake	36	0.2 – 52.6	7.53	6.0	1 - 23
Sandy Creek	32	1 - 66	32.5	10.0	10* – 170
Lund	68	0.17-12.14	5.72	3.95	0.7-31.5
Belgrade	23	0.54-18.75	7.88	5	2.5-20.9

* due to the 1mm tip size at Sandy Ck 10 mm/hr is the minimum possible value

**38 events from Narre Warren had reliable rainfall data although less were collected with pollutant samples

***based on 6-min totals

It should be noted that there are small rainfall totals in three separate catchments that are classed as events (Glen Waverley, Blackburn Lake and Lund). While they rainfall totals recorded are low in each case they triggered flow recording (using three different monitoring methods) and therefore are included in the analysis. The events recorded were also compared to the long-term rainfall records for the two sites for which long term data were available, and are presented in terms of the intensity – duration plots (Figures 4.1 and 4.2). The plots give an indication of how the captured events compare to long term rainfall records. The majority of events, as expected, fall below the 1 year ARI event. In both cases however there are at least three events above this standard.

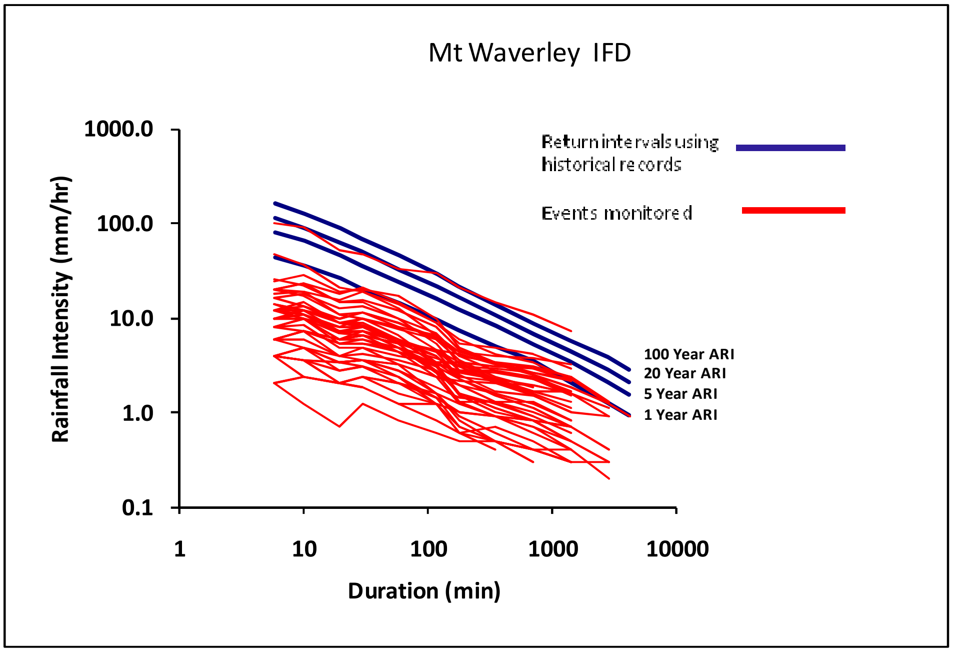


Figure 4-1 Comparison of all events and long term rainfall records at Mt.Waverley

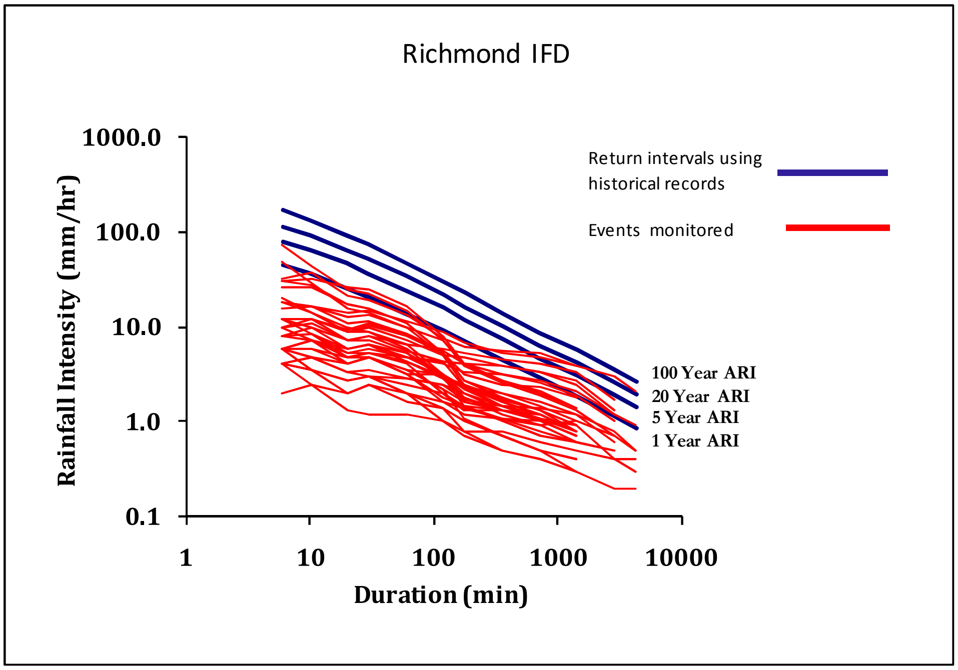


Figure 4-2 Comparison of all events and long term rainfall records at Richmond

Runoff

One purpose of the monitoring program was to describe the behaviour of a wide range of catchment sizes and land uses. To demonstrate the range of hydraulic conditions observed in the catchments, event flow data from each of these sites is presented in terms of the mean and range of the maximum flow rates for each event (Table 4-2).

Not surprisingly, the larger catchments showed greater (absolute) flow volumes. The low maximum volume recorded at Burwood East is reflective of the limited number of events, hence the lesser spread of large events, captured at the site.

The results reflect both the size of the catchment and the event size. When the mean maximum flow rates are expressed in mm/hr, the rates in part reflect the imperviousness of the catchment, as well as the rainfall intensity. Lund, Belgrade and the Roof all have high fraction impervious and corresponding high runoff rate. Narre Warren has a fraction impervious of 0.2 and a low mean maximum runoff rate.

It must be noted that the runoff is measured at the site outlet, which may be some distance from where the rainfall falls (see Chapter 6 for discussion of lag times). In these cases the travel time of the runoff will affect the results.

Table 4-2 Summary of runoff data for all sites

Site	Area (ha)	Number of events	Range of event total volumes (m^3/m^2)	Mean maximum runoff rate (l/s)	Mean maximum runoff rate (mm/hr)	Range of event maximum runoff rates (l/s)	R^2 of rainfall/runoff relationship
	Primary Data Sets						
Mt. Waverley	28.2	49	0.57 -19	408	5.2	75 - 2241	0.90
Glen Waverley	38.0	40	0.17 – 20	214	2.0	29 - 1200	0.60
Narre Warren	10.5	38	0.40 – 19	44	1.5	14 -90	0.95
Monash Roof	0.046	30	0.89 – 11.5	1.45	10.4	0.16 – 4.1	0.97
Burwood East	186	31	0.37 – 2.1	787	1.5	88.7 - 2950	0.65*
Richmond	19.2	40	0.39 – 15	547	2.2	67 -3867	0.96
Doncaster	105.7	54	0.39 – 35	723	2.5	164 - 3069	0.99
	Secondary Data Sets						
Blackburn Lake	221	36	0.11 – 27	1480	2.4	138 - 7304	0.95
Sandy Creek	227	32	0.03 – 42	4184	6.6	16 – 23,859	0.94
Lund	0.027	68	3.70 – 40	0.97	12.9	0.11 – 5.62	0.93
Belgrade	0.021	23	2.19 – 83	1.04	17.8	0.08 – 4.53	0.88

* A combination of both 6-minute offsite and 1-minute on-site rainfall were used to derive this relationship. As the total rainfall is used to derive the relationship the differing timesteps should have minimal impact.

Rainfall/Runoff relationships for all of the Primary Data sets except Glen Waverley and Burwood East show R^2 correlations greater than 0.9 (Table 4.2). Glen Waverley and Burwood East have correlations of 0.60 and 0.65 respectively which reflects difficulties in the monitoring at these sites, as discussed in Chapter 3.

Plots such as these have (see **Figure 4-3**) been used by some authors to attempt to determine at what point pervious area runoff begins (Boyd et al. 1994) by visually determining where a “break in slope” (an obvious change in the slope of the rainfall runoff plot) occurs. Previous work done on the Blackburn Lake dataset by Vaze and Chiew (2003) derived a threshold rainfall of 16mm to separate “small” and “large” events. Pervious area runoff may have different constituents and influencing process to that from impervious area runoff, and therefore an attempt was made to determine a “pervious area” runoff threshold for each catchment. Inspection of the plots however, does not indicate a clear “break of slope” in most cases.

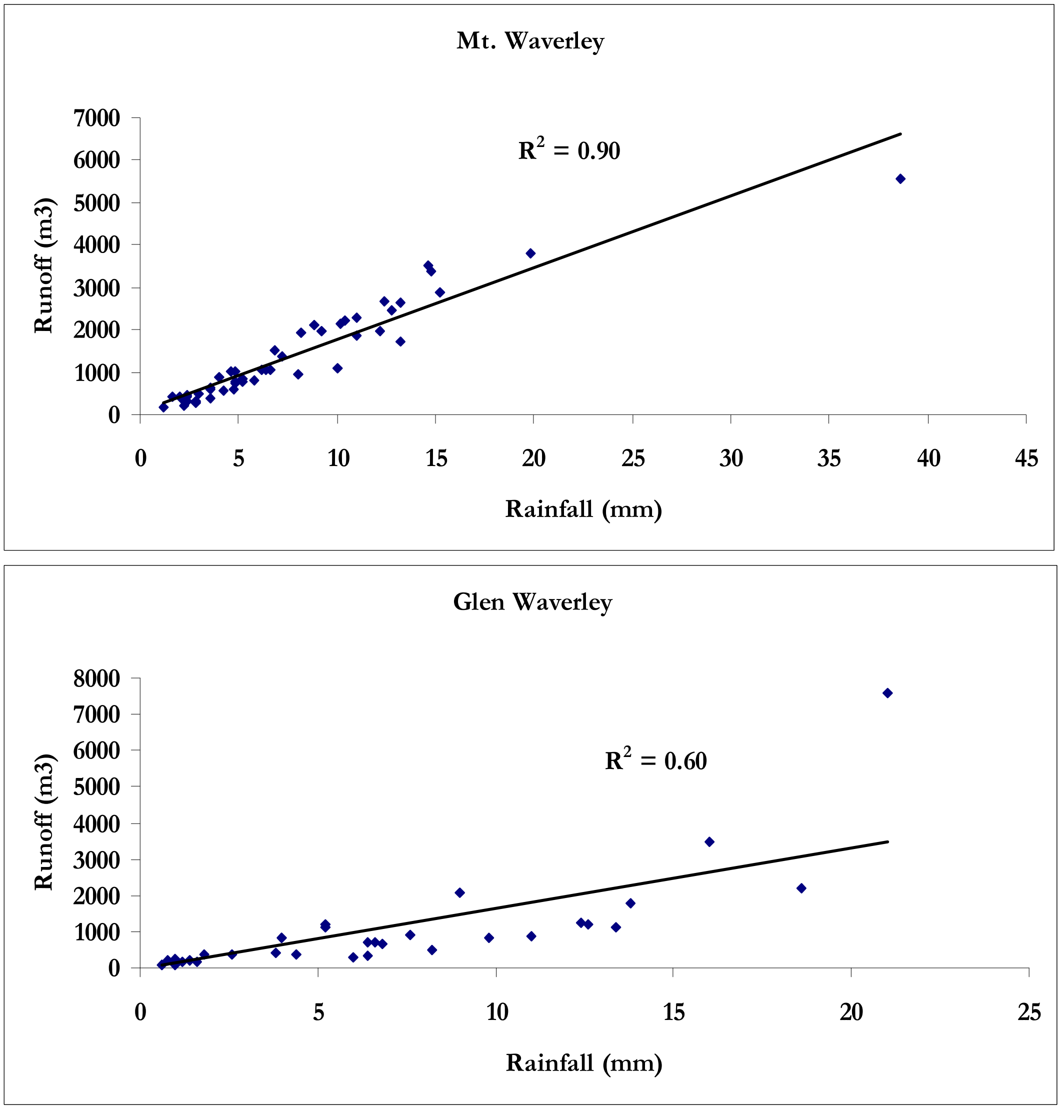


Figure 4 -3 Rainfall Runoff Plots for primary data sources

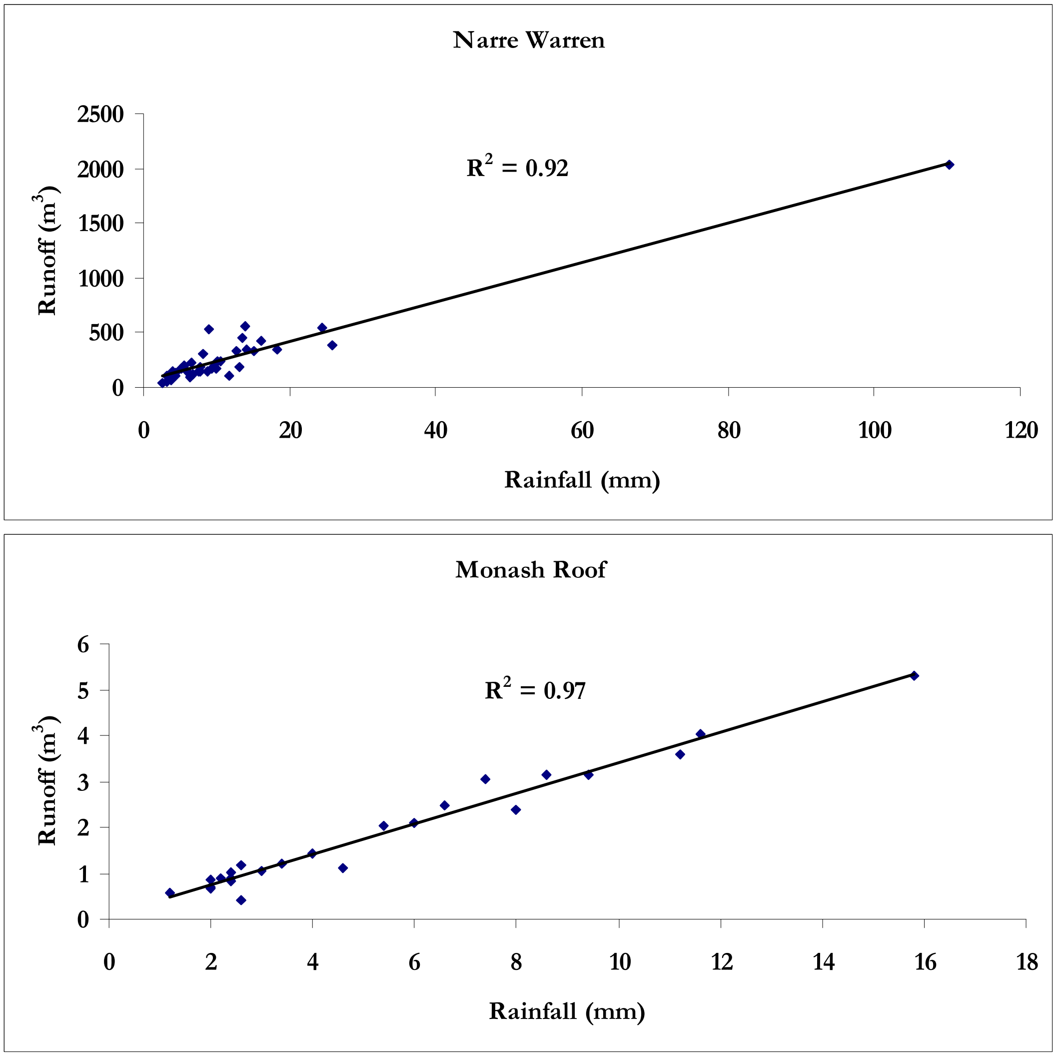
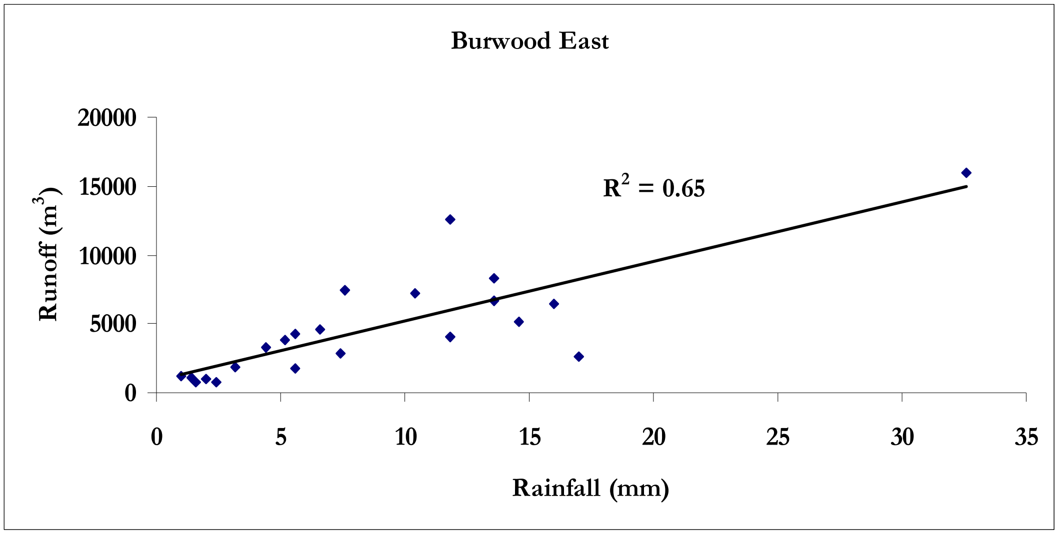
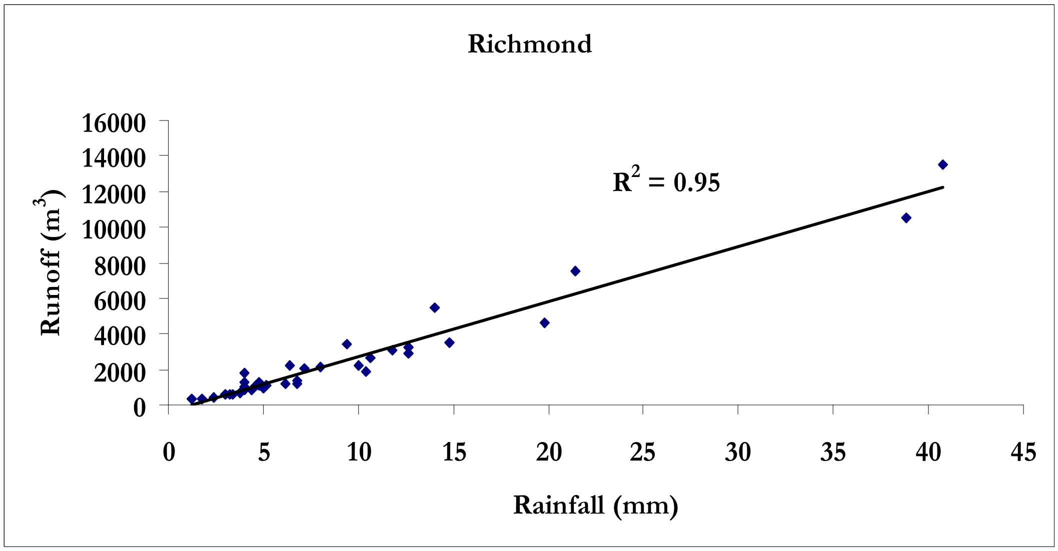


Figure 4 -3 **Rainfall Runoff Plots for primary data sources**





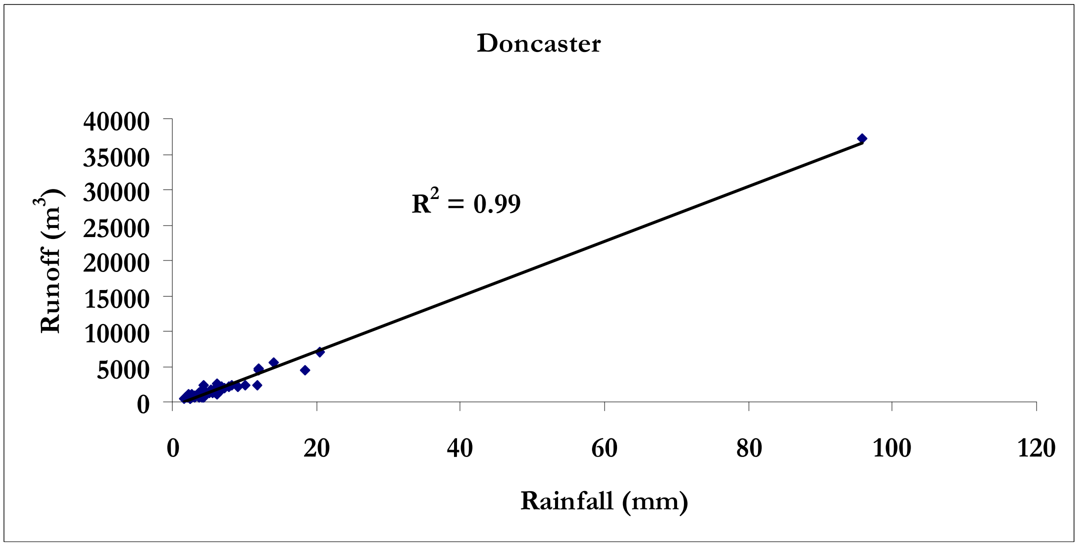


Figure 4-3 Rainfall Runoff Plots for primary data sources

It is possible that a break of slope in this case is “camouflaged” by the lack of points between 15 and 30mm of rain. Most sites have a lack of definition in this range; however few catchments show a clear break of slope even when there are data points in this range (figures 4.3 and 4.4). Rainfall intensity and seasonal effects (because of soil moisture) also will be expected to influence on the runoff coefficients, therefore it is not surprising that the “break of slope” is not clear in some cases.

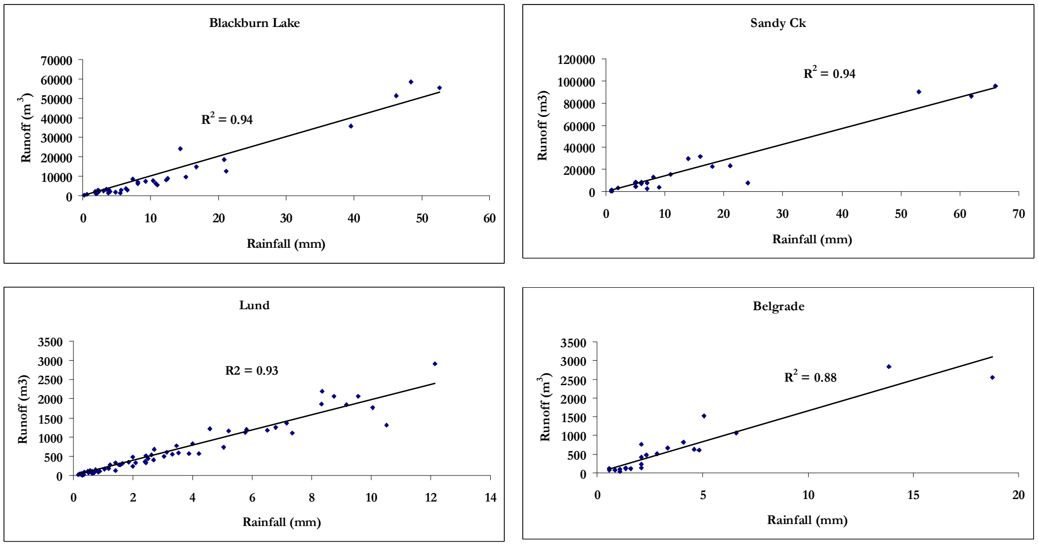


Figure 4-4 Rainfall Runoff plots for secondary data sources

Several catchments have one or more significant events which dominate the results. Narre Warren and Doncaster, in particular, captured the event from February 2005 described in Section 3-9. In both cases all events are less than 25mm in total rainfall except for the February event which totalled approximately 100mm. With this event removed the coefficient of determination drops to 0.57 and 0.86 for Narre Warren and Doncaster respectively. The Narre Warren site is the least impervious of the catchments (total fraction impervious of 0.2), and therefore more variation in runoff coefficient may be expected (Boyd et al 1994).

The Glen Waverley rainfall runoff relationship is also heavily influenced by one point (event 42 on the 2 November 2004) when 21mm of rain fell. The runoff co-efficient for this event was 0.95 (when using Equation 3.1 in Chapter 3), compared to a mean of 0.34 for the catchment. This discrepancy may indicate problems with the flow or rainfall measurement. The event was included in the analysis, although the runoff coefficient for the event suggests that one or more of the factors outlined in Section 3-9 may have been contributing. The correlation without this event is 0.67, still well below that achieved at all other sites.

Pollutants

Descriptions for the pollutant data collected are broken up into several sections. Firstly the EMC data is presented, and the raw results compared to dry weather concentrations. The SMC results are also presented and compared to catchment characteristics. The distribution of the EMC data is then confirmed, before analysis of correlations between pollutants and hydrological descriptors. Finally the loads of pollutants are presented, and then tested against the rainfall and runoff functions.

Distribution of Wet Weather EMCs

As expected from previous studies the EMC data is non-normally distributed, based on distribution-testing using the Shapiro-Wilk test (Table 4-3and Table 4-4); cases where the significance (p) of the Shapiro-Wilk test is less than 0.05 are non-normal. The log-transformed data satisfy the assumptions of normality in most cases. The exception to this is the case of TP at Burwood East, from which only 16 events were sampled. At this site, transformation actually accentuates the departure from normality. However, given the overall results, it is clear that most of the data are log-normally distributed, and the analysis described below is undertaken with log-transformed data.

Table 4-3– Normality testing for Primary data sets

Catchment	Pollutant	**Shapiro-Wilk test for normality (untransformed)**		**Shapiro-Wilk test for normality (log-transformed)**	
		Test statistic	p-value	Test statistic	p-value
Mt. Waverley	TSS	0.79	<0.01	0.98	0.64
	TP	0.70	<0.01	0.96	0.21
	TN	0.66	<0.01	0.94	0.02
Glen Waverley	TSS	0.82	<0.01	0.98	0.66
	TP	0.80	<0.01	0.97	0.49
	TN	0.80	<0.01	0.94	0.04
Monash Roof	TSS	0.47	<0.01	0.97	0.58
	TP	0.61	<0.01	0.97	0.45
	TN	0.78	<0.01	0.98	0.90
Narre Warren	TSS	0.80	<0.01	0.95	0.09
	TP	0.89	0.04	0.97	0.86
	TN	0.81	<0.01	0.92	0.15
Burwood East	TSS	0.63	<0.01	0.95	0.12
	TP	0.95	0.44	0.88	0.04
	TN	0.80	<0.01	0.89	0.06
Richmond	TSS	0.85	<0.01	0.98	0.75
	TP	0.78	<0.01	0.96	0.21
	TN	0.82	<0.01	0.96	0.16
Doncaster	TSS	0.86	<0.01	0.98	0.54

Table 4-4– Normality testing for secondary data sets

Catchment	Pollutant	**Shapiro-Wilk test for normality (untransformed)**		**Shapiro-Wilk test for normality (log-transformed)**	
		Test statistic	p-value	Test statistic	p-value
Blackburn Lake (Melbourne)	TSS	0.74	<0.01	0.98	0.70
	TP	0.76	<0.01	0.96	0.16
	TN	0.74	<0.01	0.93	0.03
Sandy Ck (Brisbane)	TSS	0.61	<0.01	0.91	0.01
	TP	0.71	<0.01	0.92	0.15
	TN	0.74	<0.01	0.91	0.11
Lund (Sweden)	TSS	0.79	<0.01	0.98	0.33
Belgrade (Serbia)	TSS	0.90	0.03	0.96	0.46

Wet Weather Events

TSS SMC and EMC values for the primary data sets are generally lower than those reported in the secondary data sets (Table 4-5). Other authors have also observed that recently collected TSS data tend to show lower EMCs than do earlier studies (Smullen et al. 1999). TP and TN results are broadly similar between the two data sets, except in the case of Narre Warren. The Narre Warren site has high concentrations of TN and TP that are probably influenced by the prevalence of septic tanks in the catchment.

The Monash Roof also has significantly lower EMC (and SMC) values than the other sites, particularly in terms of TSS and TP, which is consistent with published literature of roof catchments (Duncan 1995). The SMC for TSS is 11mg/l at the Roof compared to 54mg/l for the next lowest site (Burwood East) and above 70mg/l for all other sites.

The composition of the mean EMC (µEMC) of nutrients is presented in Table 4-6. Three of the sites have a relatively high percentage of TN in the dissolved form – Narre Warren, the Monash Roof and Burwood East. This is expected at Narren Warren due to the septic influences on the catchment runoff. The Roof was also expected to have a high proportion of nitrogen in the dissolved form due to the anticipated low concentrations of solids in roof runoff. Burwood East, however, is a typical suburban residential catchment. The high percentage of dissolved nitrogen corresponds to a high base flow concentration of TN, possibly indicating a contamination source in the catchment, or potential influences of interflow after storm events, resulting in leaching of soil nitrogen (Taylor et al. 2005).

In terms of metals far fewer events were captured, and therefore interpretation is more difficult (

Table 4-7). The two larger urban catchments Mt. Waverley and Richmond are comparable for most parameters, with the Roof site having significantly smaller EMC values for all metals, despite being a metal roof. Zinc concentrations are significantly higher than those reported elsewhere (see Chapter 2) for the two larger sites. Lead concentrations are intermediate between the published NURP data and the recently compiled US data (Chapter 2). Metals fractions are generally similar for the Mt. Waverley and Richmond catchments. The Roof site, however, had higher fractions of dissolved Copper and Zinc (51% compared to 27% at Mt. Waverley and 20% at Richmond for Copper, 90% compared to 56% and 37% at Mt. Waverley and Richmond, respectively, for Zinc). This probably reflects the smaller amount of solids available for metals transport on the roof site and less time – because of the small catchment – to adsorb onto solids. It may also reflect a direct source of metals, the roof material.

There is little evidence for land use (in terms of zoning) having an impact on pollutant loads. The only industrial catchment, Mt. Waverley, has lower SMC values than the residential catchments. The data presented has limitations however, the industrial catchment, Mt. Waverley, comprises mainly distribution warehouses with very little heavy industrial practice or manufacturing. The Richmond catchment, in comparison is high density residential; however it has a history of manufacturing and heavy industry in some parts.

Table 4-5 Summary of Event SMCs and EMCs

Site	Pollutant	Number of events	SMC (mg/l)	μEMC (mg/l)	Standard deviation of EMC's (σEMC, mg/l)
Primary Data Sets					
Mt. Waverley	TSS	49	71.4	71.6	59.2
	TP	47	0.15	0.17	0.13
	TN	47	1.02	1.17	0.82
Glen Waverley	TSS	40	76.6	94.8	79.5
	TP	36	0.20	0.24	0.18
	TN	36	1.48	1.74	0.97
Narre Warren	TSS	41	79.6	91.9	65.8
	TP	17	0.66	0.75	0.20
	TN	17	3.56	3.51	0.72
Monash roof, Clayton	TSS	30	11.0	16.7	35.0
	TP	30	0.02	0.03	0.04
	TN	30	0.44	0.53	0.69
Burwood East	TSS	32	54.0	84.1	93.2
	TP	16	0.16	0.15	0.06
	TN	16	1.58	1.54	0.54
Richmond	TSS	40	101.8	125.1	92.1
	TP	39	0.35	0.42	0.31
	TN	39	1.95	2.29	1.27
Doncaster	TSS	52	78.1	77.0	56.3
Secondary Data Sets					
Blackburn Lake	TSS	36	156	175	122
	TP	36	0.185	0.21	0.12
	TN	36	1.46	1.73	0.85
Sandy Creek	TSS	32	193	244	429
	TP	17	0.301	0.40	0.14
	TN	17	1.45	2.76	1.97
Lund	TSS	68	131	79	73
Belgrade	TSS	23	259	365	155

Table 4-6 Speciation of Nitrogen and Phosphorus

	% of TN that is					**% of TP that is**
Catchment	NH_3	NOx	TDN	DON	PON	FRP
Mt. Waverley	4%	29%	59%	25%	41%	12%
Glen Waverley	6%	32%	58%	20%	42%	25%
Narre Warren	21%	41%	79%	17%	21%	58%
Monash Roof	27%	27%	71%	18%	29%	25%
Burwood East	5%	51%	77%	21%	23%	37%
Richmond	8%	28%	57%	20%	43%	20%

*Where NH_3 is Ammonium, NOx is the combined nitrogen oxides, TDN is total dissolved nitrogen, PON is particulate organic nitrogen and FRP is filterable reactive phosphorus

Table 4-7 Metals and miscellaneous other components sampled

	Roof			**Mt. Waverley**			**Richmond**		
Parameter	Number of events	µEMC	σ EMC	Number of events	µEMC	σ EMC	Number of events	µEMC	σ EMC
pH	17	5.7	0.42	18	6.6	0.25	13	6.7	0.24
EC (uS/cm)	17	17.2	16.80	18	57.3	17.41	13	102.6	37.56
alkalinity (mg/l)	17	3.5	2.82	18	14	4.45	13	26.6	7.95
Ca (mg/l)	17	0.8	0.92	16	4.8	1.07	13	8.2	1.90
Mg (mg/l)	17	0.2	0.22	16	0.9	0.28	13	1.4	0.50
Na (mg/l)	17	1.8	1.76	16	4.9	1.25	13	7.2	2.70
K (mg/l)	17	0.2	0.12	16	1.1	0.54	13	2.5	0.98
Cl (mg/l)	17	2.4	3.56	16	5.1	2.40	13	8.9	5.85
SO4 (mg/l)	17	1.5	1.00	16	4.5	1.97	13	6.4	2.23
TOC (mg/l-C)	17	3.1	1.81	16	8.2	4.21	13	11.7	5.81
DOC (mg/l-C)	14	2.4	0.94	16	5.3	2.53	12	6.9	2.78
Al (mg/l)	12	0.89	1.17	11	7.98	14.76	8	5.30	3.15
Cd (µg/l)				10	1.3	2.6	8	0.5	0.3
Cr (µg/l)	16	3.2	4.8	18	17.3	32.2	13	13.2	6.6
Cu (µg/l)	17	6.3	6.2	18	46.3	60.7	13	56.1	25.5
Fe (mg/l)	17	1.12	1.95	18	7.40	15.10	13	6.70	3.90
Mn (µg/l)	17	21.0	26.6	18	113.5	236.8	13	116.9	54.7
Ni (µg/l)	12	2.9	3.2	11	24.0	44.3	8	16.8	6.5
Pb (µg/l)	15	5.8	6.0	18	43.3	78.4	13	112.2	60.8
Zn (mg/l)	17	0.12	0.07	18	0.95	0.79	13	1.24	0.43

Pollutant concentrations: dry weather/baseflow

Dry weather flow occurred in all catchments except the Monash Roof. This is somewhat surprising in the case of the smallest catchment (Narre Warren). With a catchment area of only 10.5 ha it would not normally be expected to flow continuously. Each house however is serviced by a septic tank and as discussed there are distinct signals of wastewater flow leaking or being incorrectly directed to the stormwater system. Narre Warren has very high concentrations of TP and TN in baseflow with median values of 7.95 and 31.5 mg/l respectively (Table 4.8).

Table 4-8 Results of dry weather sampling

Site	Pollutant	Number of samples	Median (mg/l)	Mean (mg/l)	σ (mg/l)
Mt. Waverley	TSS	14	6.75	7.65	4.79
	TP	14	0.11	0.22	0.21
	TN	14	0.92	1.13	0.62
Glen Waverley	TSS	16	14.0	20.6	20.2
	TP	16	0.24	0.23	0.12
	TN	16	2.05	2.34	1.21
Narre Warren	TSS	16	8.45	10.0	6.01
	TP	16	7.95	9.01	4.45
	TN	16	31.5	32.6	9.52
Burwood East	TSS	16	5.55	7.27	5.99
	TP	16	0.37	0.63	0.84
	TN	16	2.25	3.41	3.70
Richmond	TSS	10	6.40	12.6	17.4
	TP	10	0.22	0.42	0.68
	TN	10	9.55	11.6	10.6
Doncaster	TSS	10	13.0	16.0	10.3
	TP	10	0.22	0.24	0.13
	TN	10	2.05	2.39	1.09

TSS concentrations in event flow are at least four times higher than that in base flow at all sites (see Figures 4.5, 4.6 and 4.7). For phosphorus the concentrations are similar in most cases, except for Narre Warren as discussed above. However, nitrogen is always higher in base flow, except in the case of Mt. Waverley where it is similar. At Narre Warren, baseflow concentrations of nitrogen are significantly higher than in wet weather, due to the suspected wastewater contribution. However, overall the data supports the theory that energy, from rainfall or flow, is needed to mobilise solids in the system, but that dissolved pollutants are less affected by kinetic energy. Nitrogen in urban runoff is largely in a dissolved form, approximately 80% on average in the Melbourne context (Taylor et al. 2005), and ranging from 57-79% in this present study. At the Primary Data sites it is found at higher concentrations at base flow than during rainfall events. Nitrogen concentrations may also be significantly affected by other inputs to the system, such as septic discharge, leaking sewers or groundwater inputs. Phosphorus has intermediate characteristics between TSS and TN in terms of particulate fraction. Concentrations are higher in baseflow in all cases but the difference is not as distinct as with the nitrogen SMC’s.

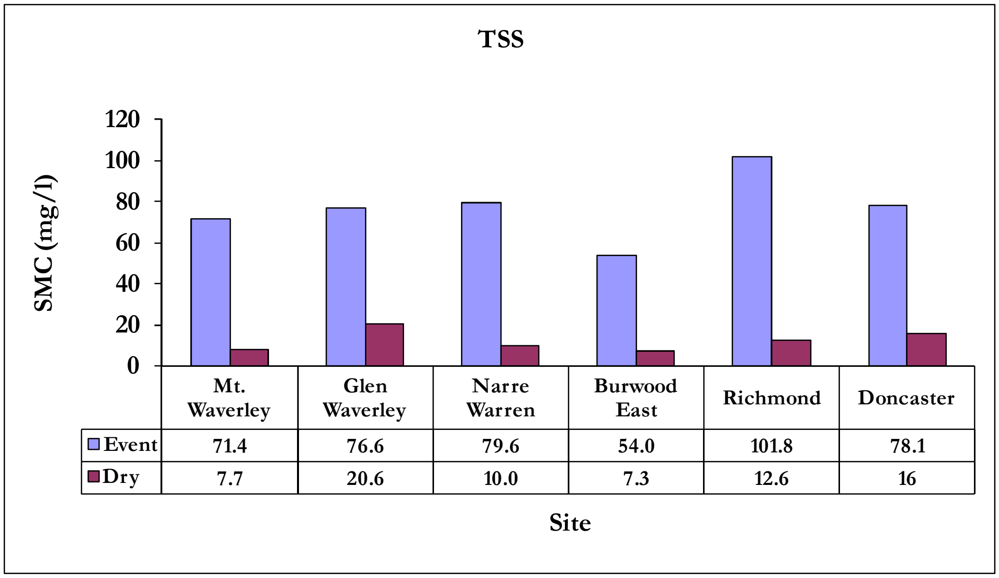


Figure 4-5 Comparison of wet and dry weather SMCs for TSS

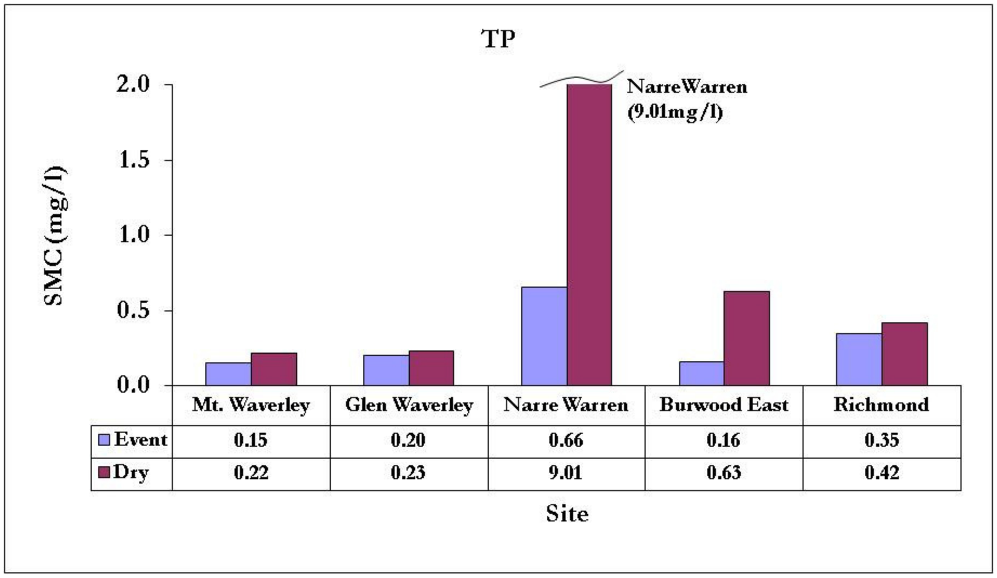
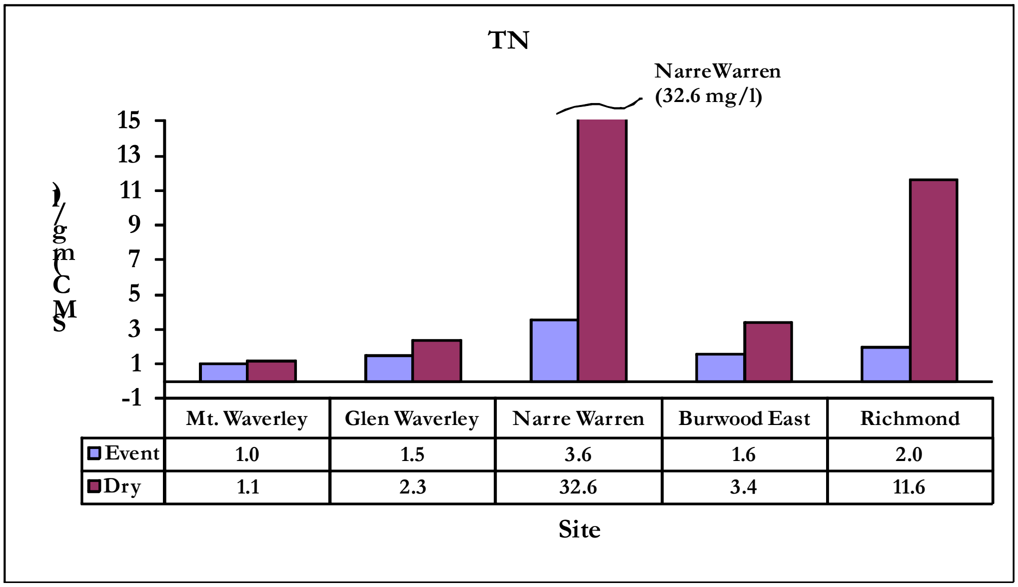


Figure 4-6 Comparison of wet and dry weather SMCs for TP



	Mt. Waverley	Glen Waverley	Narre Warren	Burwood East	Richmond
▨ Event	1.0	1.5	3.6	1.6	2.0
■ Dry	1.1	2.3	32.6	3.4	11.6

Figure 4-7 Comparison of wet and dry weather SMCs for TN

To further examine possible reasons for the differences between SMCs the approximate age of the underground infrastructure in each catchment was obtained by consulting water supply and drainage plans from each area (Table 4-6). A possible hypothesis is that as a catchment ages the underground drains deteriorate, therefore allowing more interaction with surrounding soils and groundwater. It is also possible that as the above ground infrastructure ages there is more redevelopment in the catchment, which in itself affects stormwater quality. Construction activity has been shown to have a major impact on pollutant loads, particularly when not appropriately controlled (Barfield et al. 1978).

The data does not show a strong relationship between wet weather or dry weather TSS and catchment age, possibly because most of the catchments are of a similar age and therefore do not provide enough spread (or sample size) to examine the hypothesis. The oldest catchment (in terms of land development) does have the highest wet weather TSS concentrations which may suggest that the link between age and event loads is at least worthy of further study. Figures are given from 2006 as this was the date when the monitoring program concluded.

TP and wet weather TN shows a similar lack of strong relationship with age. A plot of dry weather TN (Figure 4-8) however does show a pattern ($R^2 = 0.97$), although it is not significant at the 0.05 level because of the small number of events (four).

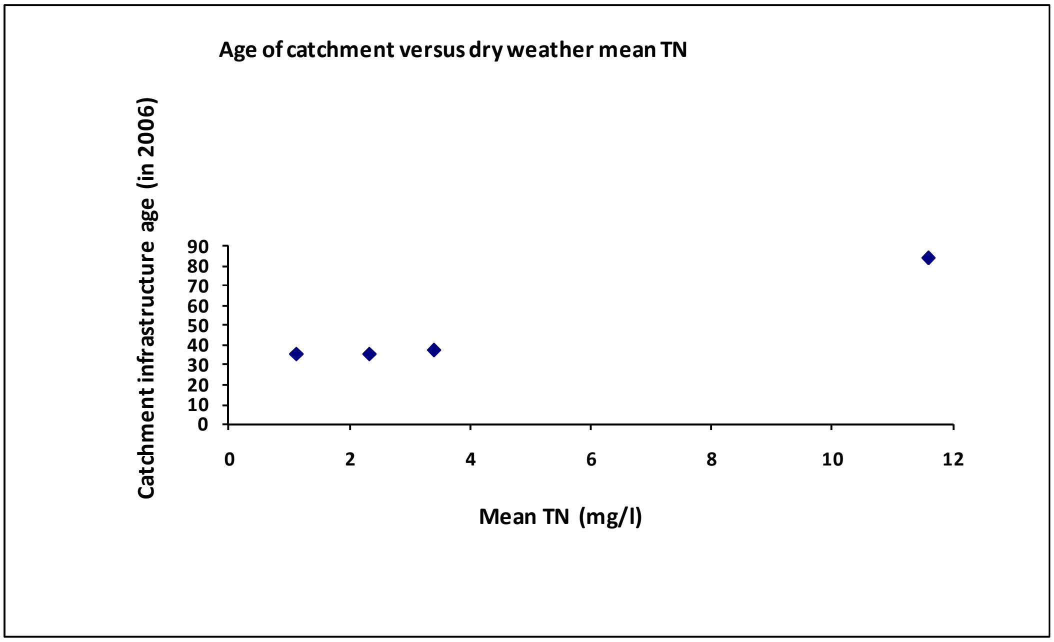


Figure 4-8 **Dry weather TN and catchment age**

As previously noted, the Narre Warren catchment was left out of the analysis for TP and TN due to the almost-certain inputs of septic system runoff.

Table 4-9 **Year in which underground infrastructure was developed within catchment**

Site	**Year of catchment development**
Mt. Waverley	1970
Glen Waverley	1970
Narre Warren Sth	1987
Monash roof, Clayton	1992
Burwood East	1968
Richmond	1922
Doncaster	1968

Correlations between EMCs and Wet Weather Event Characteristics

Correlations between the wet weather EMCs and a range of statistics representing the storm event characteristics (e.g. runoff volume, event rainfall, average intensity, etc) are examined, to determine which factors have most influence on pollutant concentrations. However, very few high correlations were recorded between the observed pollutant concentrations and rainfall, flow or antecedent conditions (Table 4-10 and Table 4-11 for the primary and secondary datasets, respectively). Maximum rainfall intensity and Maximum flow rate show reasonable correlations with TSS EMCs, however the correlations are significant at the 0.05 level in less than half the cases. Concentrations of TP and TN are less well described by the storm event characteristics. Average rainfall intensity is a poor predictor for all pollutants. Antecedent conditions (as measured by number of dry days prior to an event) are not significantly correlated with any pollutant EMCs.

Table 4-10 Correlations between explanatory variables and EMCs, primary data sets

Catchment		Correlation (Pearson's R^2/Spearman's Rho) of EMCs with storm event characteristics					
	Pollutant	**Total event runoff**	**Total event rainfall**	**Maximum rainfall intensity**	**Average rainfall intensity**	**Maximum flow rate**	**No Days < 1mm rain**
Mt. Waverley	TSS	.00/.02	.02/.06	.31/.66*	.11/.26	.33/.50*	.00/.01
	TP	.00/-.02	.01/-.09	.12*/.50*	.00/-.09	.09*/.26	.01/-.02
	TN	.05/-.09	.01/-.02	.11*/.50*	.01/-.09	.04/.26	.00/.04
Glen Waverley	TSS	.02/-.02	.00/-.01	.06/.27	.03/.13	.03/.44*	.05/.11
	TP	.07/-.26	.02/-.19	.11/.29	.00/.07	.07/.26	.13/.24
	TN	.10/-.44*	.02/-.26	.06/.02	.03/.10	.05/.18	.09/.22
Narre Warren	TSS	.02/-.24	.00/.05	.34*.62*	.27*/.52*	.18*/.53*	.01/-.06
	TP	.33*/-.77	.33*/-.86*	.00/-.07	.26/-.39	.01/-.14	.00/.00
	TN	.05/-.31	.07/-.59*	.04/-.13	.19/-.56*	.00/-.29	.01/.02
Monash Roof	TSS	.02/-.28	.00/-.15	.13/.46*	.00/.20	.16*/.53*	.17/.24
	TP	.01/-.25	.00/-.07	.04/.22	.00/.11	.05/.31	.06/.14
	TN	.26*/-.66*	.18*/-.47*	.04/-.26	.07/-.16	.02/-.17	.01/.02
Richmond	TSS	.04/-.10	.04/-.09	.19*/.45*	.04/.13	.11*/.28	.10/.16
	TP	.05/-.15	.04/-.09	.11/.34*	.02/.18	.05/.19	.04/.12
	TN	.08/-.26	.06/.20	.04/.18	.00/.12	.01/.02	.03/.08
Burwood East	TSS	.08/-.33	.02/-.11	.06/.39	.13/.16	.21/.16	.19/.40
	TP	.05/.30	.04/.38	.03/.31	.08/.76*	.07/.51	.05/.24
	TN	.02/-.01	.04/.19	.20/.40	.24/.59	.02/.27	.04/.03
Doncaster	TSS	.00/-.04	.01/.17	.06/.28	.28*/.52*	.11*/.44*	.04/.03

* significant at p = 0.05

Table 4-11 Correlations between explanatory variables and EMCs, secondary data sets

Catchment		Correlation (R^2/Spearmans Rho) with log transformed data					
	Pollutant	**Total event runoff**	**Total event rainfall**	**Maximum rainfall intensity**	**Average rainfall intensity**	**Maximum flow rate**	**No Days < 1mm rain**
Blackburn Lake	TSS	.00/.09	.00/.09	.40*/.54*	.01/.16	.26/.38*	.01/.06
	TP	.03/-.15	.03/-.14	.21/.28	.00/-.02	.13/.18	.04/.22
	TN	.08/-.35*	.05/-.37	.02/.12	.06/-.22	.01/-.11	.04/.33
Sandy Ck	TSS	.01/.28	.02/.33	.06/.53*	.09/.53*	.03/.52*	.00/.13
	TP	.00/.08	.00/.18	.04/.56*	.02/-.31	.00/.12	.01/-.09
	TN	.14/-.28	.17/-.35	.56*/.49*	.28*/-.66*	.13/-.28	.03/.35
Lund (Sweden)	TSS	.09/.30*	.12/.34*	.00/-.01	.00/.07	.00/.07	.00/.05
Belgrade (Serbia)	TSS	.17*/-.67*	.15/-.49	.01/-.35	.01/-.45*	.05/-.57	.03/.04

* significant at p = 0.05

The log EMC values are compared between pollutants to assess any cross correlation (Table 4-12). Correlations between TSS and TP, and TSS and metals have been reported frequently in the literature including for Australian datasets (Sartor and Boyd 1972; Sharpin 1995).

There is usually some correlation between different pollutants at a particular site. TSS and TP are strongly correlated at most sites, the exception being Narre Warren which can again be explained by the influence of the septic systems in the catchment. Burwood East has been left out of this discussion due to the small number of events sampled for TP and TN. For most sites the relationships between TSS-TP and TP-TN are stronger than those for TSS-TN. This is consistent with previous studies which have shown that phosphorus can occur in particulate form, and is adsorbed onto particulates in urban runoff (Cowen and Lee 1976; Randall et al. 1982; Hvitved-Jacobsen et al. 1984). It is therefore expected to be strongly correlated to TSS. Data from a previous study on the Doncaster catchment showed that approximately 75% of the total phosphorus was in particulate form (Taylor et al. 2005). On the other hand, the majority of nitrogen is in the dissolved form (Table 4.4), and is thus less likely to show correlations with sediment concentrations. At Narre Warren TP/TN are strongly correlated, again suggesting that contamination from septic systems is the source of the high concentrations.

Table 4-12 Cross correlations between pollutants (primary data sets)

Correlation (R^2/Spearmans Rho) for log transformed EMCs					
Pollutants	**Mt. Waverley**	**Glen Waverley**	**Narre Warren**	**Monash Roof**	**Richmond**
TSS - TP	.68*/.81*	.72*/.84*	.03/.21	.55*/.70*	.68*/.89*
TSS - TN	.20*/.48*	.47*/.69*	.00/-.10	.36*/.31*	.82*/.78*
TP - TN	.55*/.61*	.74*/.85*	.48*/.64*	.37*/.50*	.86*/.94*

* significant at the 0.05 level

Table 4-13 Cross correlations between pollutants (secondary data sets)

Correlation (R^2/Spearmans Rho) for log transformed EMCs		
Pollutants	**Blackburn Lake**	**Sandy Ck**
TSS - TP	.79*/.86*	.27*/.56*
TSS - TN	.16*/.61*	.56*/.49*
TP - TN	.26*/.71*	.06/.25

* significant at the 0.05 level

Relationships between event loads and explanatory variables

This section builds on previous work which attempts to relate hydrological and catchment variables to pollutant loads (Chiew et al. 1997; Vaze and Chiew 2003). It is also the first step in testing the central hypothesis of this thesis (see Equation 4.2).

The optimised rainfall and runoff functions (Section 4-3), the total rainfall and the total runoff volume are compared to the total event loads (Table 4-14). Note that the total rainfall and total runoff volume are a special case of the rainfall and runoff functions, when b and d each equal 1, however they are displayed separately to enable analysis of the equation form (there is no value in having the b or d terms if they are always equal to 1). Several points can be noted:

- Correlations are generally high as may be expected, as part of the correlation relates to the total size of the event, and all functions tested are to some degree correlated with event size. This behaviour will be further explored in Chapter 5.

- The rainfall function generally gives the highest correlations for all pollutants, particularly when the primary data sets are considered.

- In the majority of cases the rainfall function gives a significantly better correlation than the total rainfall or total runoff volume. Importantly this indicates that the power (b in equation 4.2) is important and the results do not solely reflect a relationship between volumes and loads.

Table 4-14 Correlation (coefficient of determination) of pollutant loads with rainfall function and rainfall volume with one-minute rainfall (maximum values in bold)

Catchment	Pollutant	Total event rainfall, ΣI	Rainfall function, ΣI^b	Total event runoff, ΣR	Runoff function, ΣR^d
Primary data sets (using one-minute rainfall data)*					
Mt. Waverley	TSS	0.78	**0.91**	0.68	0.88
	TP	0.88	**0.89**	0.87	**0.89**
	TN	0.92	**0.94**	0.85	0.91
Glen Waverley	TSS	**0.83**	**0.83**	0.65	0.73
	TP	**0.81**	**0.81**	0.67	0.73
	TN	**0.89**	**0.89**	0.77	0.85
Narre Warren	TSS	0.88	**0.91**	0.82	0.87
	TP	0.87	0.89	**0.93**	**0.93**
	TN	0.94	0.94	**0.98**	**0.98**
Monash Roof	TSS	0.60	**0.62**	0.38	0.46
	TP	0.59	**0.60**	0.56	0.58
	TN	**0.76**	**0.76**	0.74	0.75
Richmond	TSS	0.73	**0.90**	0.68	0.73
	TP	0.82	**0.90**	0.79	0.79
	TN	0.91	**0.94**	0.89	0.89
Burwood East	TSS	0.93	0.93	**0.94**	**0.94**
Doncaster	TSS	0.94	**0.96**	0.93	0.93
Secondary data sets (using one-minute data except in the case of Blackburn Lake)*					
Blackburn Lake	TSS	0.80	**0.94**	0.84	0.92
	TP	0.84	0.85	0.89	**0.90**
	TN	0.89	0.90	0.96	**0.97**
Sandy Creek	TSS	0.82	**0.87**	0.83	**0.87**
	TP	0.95	**0.96**	**0.96**	**0.96**
	TN	**0.98**	**0.98**	0.94	0.95
Lund	TSS	0.75	**0.76**	0.72	0.72
Belgrade	TSS	0.72	0.72	0.86	**0.87**

* All regressions significant at 0.05 level - bold designates the strongest correlation(s) for each case

The Monash Roof shows the weakest relationship with both the rainfall and runoff functions. A plot of the Rainfall function vs the TSS load at the Roof site (Figure 4.9) demonstrates that this is largely caused by two outlying points. These points represent events where the EMC value is significantly higher than others measured at the site (EMCs of 93 and 190 mg/l compared to an average of 11mg/l). These points may be a function of catchment characteristics (i.e. small catchments may be more susceptible to variation due to unusual events). Alternatively the outliers could reflect an inconsistency in the monitoring procedure such as material building up on the probe. For the purposes of model development, the analysis has been repeated without the two outlying events (Table 4.14).

Events 18 and 23

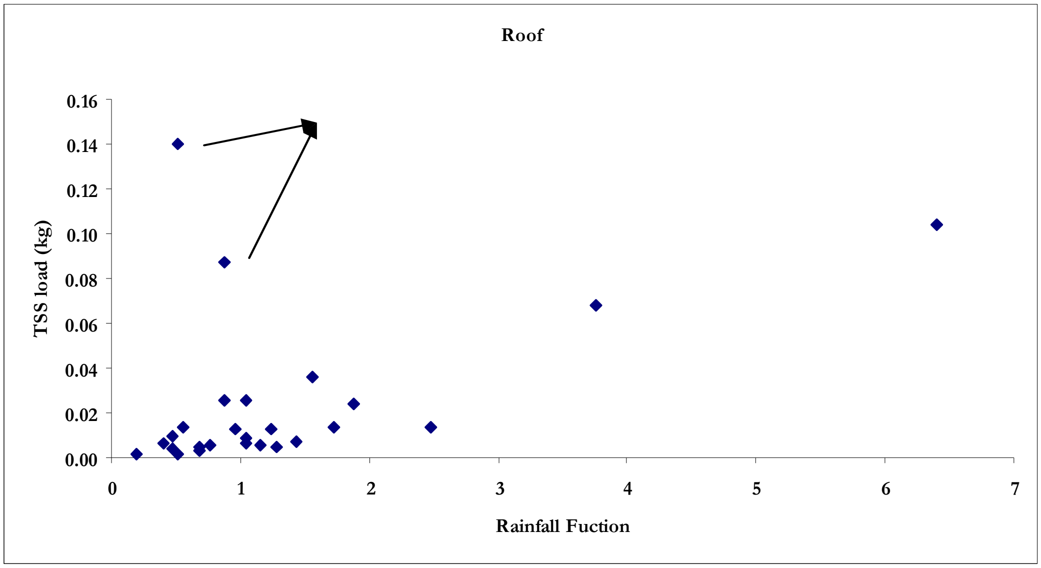


Figure 4-9 Rainfall Function vs TSS load at the Monash Roof site showing the two outlying events

Table 4-15 Monash Roof correlations without outlying events (18 and 23)

Catchment	Pollutant	Total event rainfall, ΣI	Rainfall function, ΣI^b	Total event runoff, ΣR	Runoff function, ΣR^d
Monash Roof (with events 18 and 23 removed)	TSS	0.67	**0.93**	0.67	0.91
	TP	0.82	**0.85**	0.78	0.81
	TN	**0.88**	**0.88**	0.85	0.87

4.5 Discussion

Rainfall

The rainfall results are generally as expected, with Lund having the lowest mean maximum rainfall intensity and Sandy Creek (Brisbane) the highest. All of the Melbourne sites are comparable.

When compared to long-term records, Mt. Waverley and Richmond show a reasonable coverage of rainfall events. In both cases several events of greater than 1 in 1 Year ARI are recorded, and in the case of Mt. Waverley one event of 1 in 100 Year ARI was captured.

Runoff

The runoff results also show that a wide range of event sizes has been captured. As expected, when the mean maximum runoff rate is expressed in mm/hr, there is a clear relationship with imperviousness of the catchment. Lund, Belgrade and the Roof have high values, and Narre Warren and Glen Waverley have low values. Richmond has lower than expected values and it is hypothesised that this could be due to a “hydraulically inefficient” (i.e. old and leaky) drainage system.

The relationship between rainfall and runoff for the catchment, which is used for preliminary evaluation of data quality, reflects the difficulties that arose during the monitoring program, as previously discussed in Chapter 4. Importantly the uncertainties created by these occurrences are considered in Chapters 5 and 7.

Pollutant concentrations

The TSS SMCs (and EMCs) are significantly lower for the Primary data sets than for the Secondary data sets and previously published results (Duncan 1999). This can partially be explained by the nature of the catchments and the points from which the samples were taken. Both Sandy Creek and Blackburn Lake catchments have sections of open waterway upstream of where the monitoring is undertaken. These sections of open waterway may contribute excess TSS (TN and TP EMCs are within the range of those events collected in 2003/05). Interestingly, a similar pattern was observed in the USA, when an update of the Nationwide Urban Runoff Program data was undertaken in the late 1990s (Smullen et al. 1999; Pitt and Maestre 2005). It is also possible that, at least in part, the trend may be a result of refined monitoring procedures such as autosampler type and sampling regime (number of samples taken to describe an event). Past reported studies in Melbourne and Brisbane (RossRakesh et al. 1999; City Design 2003) have recorded TSS EMCs similar to those reported in Fuchs et al. (2004). These studies, however, have been recorded from catchments which included sections of natural waterway. It is possible that these sections contribute TSS through in-stream processes (sediment scour and bank slumping, for example). Adding support to this argument, the TN and TP concentrations have not shown comparable differences. Another contributing factor may be the recorded reduction in atmospheric pollution, which is resulting in cleaner cities and therefore cleaner runoff. Substantial reductions of concentration in air of many trace elements have been reported over the period of interest (Lee et al. 1994; Var et al. 2000). The finer particles of health-related interest (PM_{10}) also show reductions, both in emissions to air (Harrison 2002) and in concentrations in air (Wang et al. 2006). However, the larger suspended particulate matter shows much less change (Jakeman et al. 1992; Var, Narita et al. 2000). Improved air quality may contribute to reduced TSS concentrations in urban runoff, but it appears unlikely to be the sole cause.

TSS values from the Monash roof have a mean of 11 mg/l which is very low compared to other sites which are in the 50 –100 mg/l range (but comparable to other studies which have compared TSS from roofs and other urban catchments (Bannerman et al. 1993, Duncan 1999). This may be because of the energy required to lift particulates onto roofs compared to road or other ‘low’ surfaces. TP values from the roof are typically an order of magnitude less than those from other urban catchments, and TN are half the value of the next smallest catchment.

There is also a greater variation between sites in the EMC values of TP and TN, when compared to TSS, with values recorded at the Narre Warren catchment significantly higher than those recorded elsewhere. When Narre Warren is excluded the TP and TN EMCs still vary over a greater range than TSS.

The comparison of Dry and Wet weather values represents an important finding with TN having significantly higher concentrations in baseflow than during storm events. Little data on the comparison between base flow and stormflow exists in Australia although some international research has reported that baseflow concentrations for many constituents are often significantly less than stormflow (McPherson et al. 2002).

The high values of baseflow TN have a number of potential sources:

- In Narre Warren they are almost certainly coming from leaking septic systems.

- In other urban catchments it is likely that aging sewer systems are contributing to base flow (and possibly storm flow) TN and TP loads. Several catchments of similar ages to the catchments monitored have been investigated for sewage leaks as part of Melbourne Water’s drainage investigations (Graham Rooney, Melbourne Water, personal communication) and have had leaks or illegal connections with the sewer system discovered.

Investigations by Taylor et al (2005,2006) have shown that TN concentrations often increase at the end of storm events, and continue to remain high for several days after an event in Melbourne. Taylor et al. (2006) hypothesised that this may be due to leaching of nitrogen from soils, occurring during interflow and throughflow. Obviously, the magnitude of such leaching will depend on the soil characteristics and the past land uses, as well as possible nitrogen fixation from plants such as Acacias (Sprent 2000).

Cross correlations between pollutants are strong at most sites (except Burwood East when only a small number of events were captured). TSS – TP correlations are usually the strongest, followed by TP – TN. This is consistent with previously published data (Sharpin 1995). At Narre Warren the correlations between TP/TN and TSS are low, supporting the theory that septic sources (which are high in nutrients but low in TSS) may be responsible for the observed results.

Pollutant loads

The addition of the power function (*b* in equation 4-2 or *d* in equation 4-3) in the majority of cases improves the correlation with event load over that of the total event runoff or rainfall models. The extent to which the power function outperforms the simple rainfall total, for either rainfall or runoff, is a useful measure of modelling effectiveness. We could not justify the inclusion of an additional coefficient (the power, *b* or *d*), unless it provides a useful improvement in predictive performance at most sites. The power function also demonstrates that the relationships obtained are not purely a spurious relationship based on the size of the event; ie. it delivers *additional* information beyond that provided simply by the total event rainfall depth or runoff volume.

This suggests that the pollutant load washed off a catchment may be determined not only by the total rainfall or flow, but also by the *input of energy* to mobilise pollutants in the catchment. The improvement gained by including a measure of 'instantaneous' rainfall intensity in most cases is largest for TSS, and rather less for TP and TN. This again supports the case for energy as a key component of the washoff process, since TSS is completely particulate while TP and TN are partly dissolved. Dissolved species do not require the same amount of energy for their mobilisation and transport. The power function is investigated further in Chapter 6.

The preference for rainfall or runoff as the underlying explanatory variable based solely on Table 4-14 is clear. Rainfall performs better in most cases for all pollutants. On this basis alone rainfall intensity is the preferred variable, particularly for TSS.

There are also good practical reasons to prefer rainfall over runoff as an explanatory variable. Rainfall records are much more widely available at the short time intervals needed for the power functions, and are not affected by local land use change.

A number of points may be noted:

- Firstly, the power *b* has been optimised by iteration for all events at each site separately. A power of 1.9 would indicate a pollutant that increases markedly in concentration as rainfall intensity increases. For comparison, a power of 1.0 describes a pollutant that maintains a constant concentration regardless of rainfall intensity. A power of $b < 1$ indicates that the concentration decreases with increasing intensity and a power of $b > 1$ indicates the concentration increases with increasing rainfall intensity.

- The screening approach based on R^2 at each site separately tells us only that a good relationship can be found at each site separately. We have not yet considered the magnitude of the relationship, or the effect of timestep length and catchment area, or the relationship between sites. Following the arguments of Vaze & Chiew (2003), we propose that total load in a timestep at a given rainfall intensity is proportional to the length of the timestep – twice the duration yields twice the load. It is also implicitly assumed that total load is proportional to total catchment area.

4.6 Conclusions

The initial analysis outlined in this chapter provides a number of useful conclusions. On a local and even national scale, this new Melbourne dataset is a first in terms of the rigorous collection methods and the number of events collected. The monitoring programme did not attain its original goal of collecting approximately 50 events (for TSS) from 8 catchments; however it does provide a high degree of detail about a number of catchments. The EMCs reported are generally less than those currently used in Melbourne (based on previous data from international studies) for TSS, TP and TN, and should provide a valuable data set for practitioners. The metals data is also the most extensive set of data currently available. The comparison of base flow and stormflow concentrations reported is in some cases contrary to current understanding and practice (with higher concentrations often during baseflows) within stormwater management in Australia and should be further investigated. The monitoring program did not accurately record flow rates for base flow at any of the sites, but, given the high concentrations of TP and TN it is possible that baseflow is a significant contributor to total loads. Based on the results from Narre Warren it is also likely that base flow concentrations of TP and TN in particular are very dependent on the state of sewage and other assets in urban areas.

The lack of strong relationships between EMCs (or SMCs) and a number of hydrological variables is consistent with literature from Australia and internationally. This further underlies the difficulty in predicting stormwater quality concentrations, a problem that has been recognised in many studies. All concentration data sets are log-normally distributed, which supports previous comparisons of urban EMC data (Duncan 1999). Correlations between EMC values and the storm event characteristic variables typically show weak or no relationships. For TSS, the maximum rainfall intensity provides significant correlations for most catchments (when the Spearman Rank method is used).

Encouragingly, the proposed rainfall and runoff function equations show good potential for predicting pollutant loads, for all data sets. They also consistently show that the rainfall function performs better than does the runoff function. This supports the original hypothesis by Duncan (1995) and the revised hypothesis presented in Chapter 3. Most importantly, the rainfall function performs better than a purely volume based approach, indicating that the pollutant load is not only a function of rainfall or flow volume, but is driven also in part by the intensity of rainfall, and thus by its kinetic energy.

5 EVENT LOADS

5.1 Introduction

Chapter 4 of this thesis concluded by using regression analysis to examine the possible drivers of pollutant loads. Correlations between total rainfall, total runoff and the runoff function were quite strong however the function of the rainfall intensity generally gave the best results for all pollutants for the majority of catchments. These encouraging results suggest that a simple model, based on widely available data (short timestep rainfall), may be able to effectively predict pollution loads from urban catchments. Further work is required however, to examine if this basic regression analysis can be transferred into a usable predictive tool.

The central questions to be addressed in this chapter are the ability of the model to predict pollutant loads for individual catchments when calibrated, and then for catchments independent of the calibration data. As discussed in Chapter 2 many models have performed well when calibrated against a particular catchment, but few are able to adequately predict the behaviour of nearby catchments, let alone unrelated catchments.

The chapter also begins to explore the structure of the model by comparing the way calibration parameters behave for the dataset. Finally, the chapter will evaluate the link between timestep and the other calibration parameters, and will propose a set of parameters to be used and further evaluated.

5.2 Background

Prediction of event loads for unmonitored catchments is the primary goal of many stormwater management programs; indeed some authors have questioned if anything beyond event loads are ever required (Huber 1992). Well known models such as SWMM (Huber and Dickinson 1988) use a runoff rate-based model with a similar formation as the proposed model, although they usually include a term for buildup and do not sum the rainfall intensity over a whole event to calibrate parameters. Calibration coefficient values for SWMM type models vary significantly (Jewell and Adrian 1982; Tsihrintzis and Hamid 1998; Calabro 2001; Choi and Ball 2002).

Calibrations for SWMM have been reported as good, for single, intensively studied catchments, if reported loads are within approximately 20% of observed loads (Baffaut and Delleur 1990; Tsihrintzis and Hamid 1998).

Equations derived from multiple regression analysis have also been used to predict events loads, in particular those resulting from large French and US studies in the 1980s. Using equations containing the maximum rainfall intensity over a five minute period and the runoff volume, the French program reported predicted loads within approximately 10% of observed long term loads using these equations, and within 30% for individual events (Desbordes and Servat 1984; Desbordes and Servat 1987). Regression models from the USA's NURP program have been reported to explain a similar variance in predicted versus observed loads (Brezonik and Stadelmann 2002).

The method proposed by Duncan (1995) and investigated by Chiew (1997) and Vaze and Chiew (2003; 2003b) proposes rainfall intensity as the key variable describing pollutant generation, but retains simplicity in the model. This model has been introduced in chapter 2 and investigated in chapter 4. Vaze and Chiew achieved the best results with this method for relatively large catchments, when a combination of rainfall and runoff was used. The addition of the extra term however, and two associated calibration coefficients, comes with a price of declining model simplicity and increasing data requirements to support the model. Problems associated with model over-parameterisation have been noted elsewhere (Ginzburg and Jensen 2004). The strength of the regression analysis presented in Table 4-14, particularly for the monitoring program designed specifically to test the theory, suggest that this price may be too high.

The model formulation discussed in chapter 4 is:

$$Load = W \times A \times T \times \sum_{i=1}^{n} I^b \qquad \textbf{Equation 5-1}$$

where *Load* is the load of pollutant produced by the whole catchment over the period of interest (units of mass), *W* is a calibration coefficient dependent on pollutant and location, *A* is assumed to be the total catchment area (units of length^2), *T* is the duration of the timestep (units of time), *I* is the rainfall intensity in the timestep (units of depth/time), *b* is a constant dependant on the pollutant, and *n* is the number of time steps over the period of interest.

The power *b* is dimensionless, but the washoff coefficient *W* has the difficult units of mass × length^{-2} × time^{-1} × (length/time)$^{-b}$. When $b = 1$ this simplifies to units of concentration, but in the general case the units of *W* depend upon the power *b*. The practical significance of this is that the washoff coefficient *W* cannot be compared between catchments unless *b* is the same in each case.

The proposed model is similar to others reported elsewhere (Desbordes and Servat 1984; Kanso et al. 2004) that have not provided consistently good results. However there are three contributing factors regarding the current program that may be contributing to the promising results reported in Chapter 4.

- Firstly, the primary dataset provides a number of sites monitored using a consistent and well controlled methodology, in a particular climatic zone and with short interval rainfall data available from the catchments in question

- Secondly, the monitoring is exploring the integrated effect of short timesteps on prediction. As discussed many previous attempts to determine the primary variables for pollutant loads have used averages rainfall intensities or large timesteps

- Thirdly, the program is designed to rigorously collect a sufficient number of events to describe the catchment behaviour to a known level of uncertainty

5.3 Method

Chapter 4 introduced the concept of rainfall and runoff functions, and showed how the rainfall intensity based function generally gave better correlations than the others tested. In this chapter the influence of the rainfall function on pollution loads is further analysed.

In general terms the model needs to now be tested with a greater degree of rigour, including against catchments not used for its calibration. The specific questions that need to be addressed to develop the model further include:

- Can the encouraging results presented in Table 4-14 provide any information that can be used to refine the model?

- With an appropriate calibration and validation procedure, can the proposed model adequately predict event loads? The results are compared against both published results from comparable models and with the 30% experimental uncertainty that was proposed in Chapter 4

- If the model can satisfy the point above, can a set of parameters be determined to allow the model to predict event loads from catchments independent of the calibration data set?

- If successful in prediction, are the calibration parameters associated with the model independent, and if not (ie. if they are co-correlated) is the model's level of complexity justified by the data available?

- What is the impact of model timestep and rainfall gauge sensitivity on model performance?

The first step taken was to compare plots of the rainfall function and observed loads to detect any patterns in the data. A series of calibration activities were then undertaken which firstly looked at each catchment in isolation, and then attempted to compare calibration across catchments. Finally the sensitivity of the model parameters was examined.

Examination of rainfall function/event load plots

To provide an overview of the data sets available, the first step in this process was to analyse the rainfall function versus load plots to further understand both the quality of the data and the nature of the relationships. The linear correlation (coefficient of regression) was calculated for all catchments. In particular, the plots were examined for signs of curvature in the data that would suggest a solution other than a linear regression.

The significance of the magnitude of the y-intercept was then examined for all catchments and pollutants. A y-intercept close to zero (or not significantly different from zero) would simplify the model by suggesting that pollutant generation only begins when enough energy (represented by the rainfall function) is supplied to mobilise pollutants. If the Y-intercept is significantly different from zero it would suggest that another term may be needed in the model (such as a buildup term) to fully explain the variation.

If the Y-intercept is not significantly different, it is proposed that the straight-line regression relationship can be further simplified to a straight line through zero.

Validation for individual catchments

The model was then tested against each individual catchment in isolation. Given the high correlations shown in Chapter 4 (Table 4-14), a more rigorous exercise was carried out by calibrating the proposed model and then producing a series of predicted event loads for each event and pollutant.

A cross-verification method was used for this analysis according to the method described in Vaze and Chiew (2003). One event in each dataset was removed and calibration parameters were calculated from all remaining events. These parameters were then used to predict the event load of the removed event. This process was in turn repeated for all events in a catchment. Therefore if there were *N* events in the data set, *N-1* events were used for calibration. The parameter set generated from the *N-1* events was applied to the excluded event, and hence an event load was calculated.

The method is demonstrated using a sample graph of TSS load vs Rainfall function for Blackburn Lake (Figure 5.1). The plot is a graphical representation of the results presented in Chapter 4 (Table 4-134) which presents the strength of the relationship between the rainfall function and the event load. Therefore in Equation 5-1 b = 1.8 and *WAT* = 327 for TSS at Blackburn Lake. These parameters are calculated using the whole data set. In contrast the cross calibration exercise requires a separate parameter set (*b* and *W*) to be calculated for all events individually by excluding the event in question from the calibration.

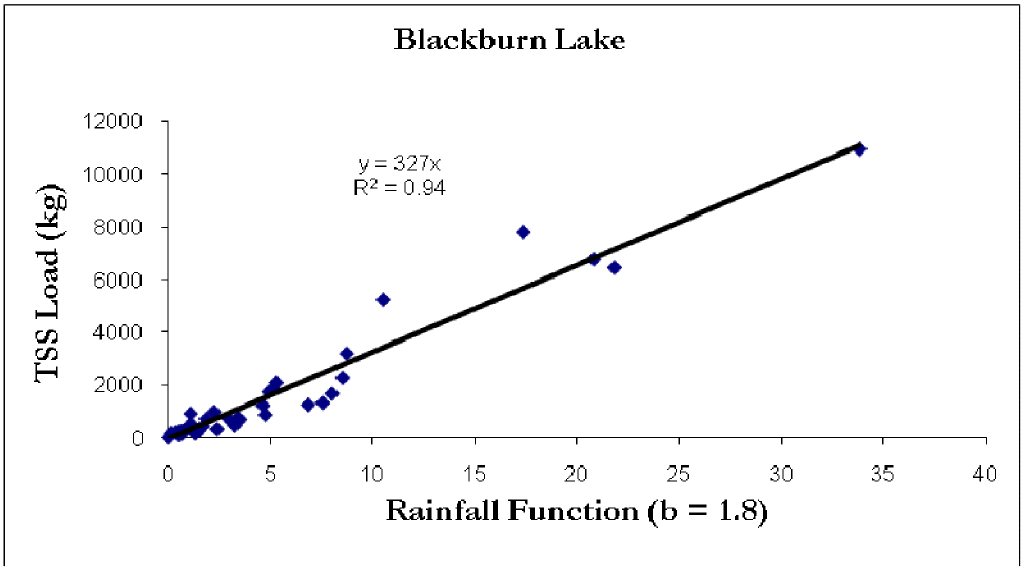


Figure 5-1 Sample results demonstrating the calibration method (the rainfall function is the sum of the intensity to the power of 1.8, as detailed in $Load = W \times A \times T \times \sum_{i=1}^{n} I^{b}$ **Equation 5-1**

The results of the verification exercise were reported against both the total load from the catchment and against individual events. Three objective functions were used. The three functions were:

- The ratio of the observed versus predicted loads for the total dataset (total pollutant load over all recorded events).
- The coefficient of determination of the observed versus predicted event loads. The R^2 value compares the magnitude of the predicted versus observed loads for each event to other events in the data set, but will not detect any systematic bias (Figure 5-2).
- The co-efficient of efficiency (E) of the observed versus predicted event loads (Nash and Sutcliffe 1970). The E value compares the predicted results to the 45 degree line on a graph, therefore picking up any systematic bias in the results.

The coefficient of efficiency is defined as:

$$E = \frac{\sum_{i=1}^{n}(REC_i - \mu REC)^2 - \sum_{i=1}^{n}(EST_i - REC_i)^2}{\sum_{i=1}^{n}(REC_i - \mu REC)^2}$$

Equation 5-2

where REC and EST are the recorded and estimated loads respectively, and µREC is the mean value of recorded loads. The coefficient of efficiency provides the proportion of variance in predicted loads explained by the model. The E value is always less than or equal to a corresponding coefficient of determination value. An E value of 0 indicates that the mean observed value is an equivalent predictor to the parameter tested (Legates and McCabe 1999).

The first of these functions gives a good method of comparing the models effectiveness at predicting long term loads, the second and third give clarity around its ability to predict individual events.

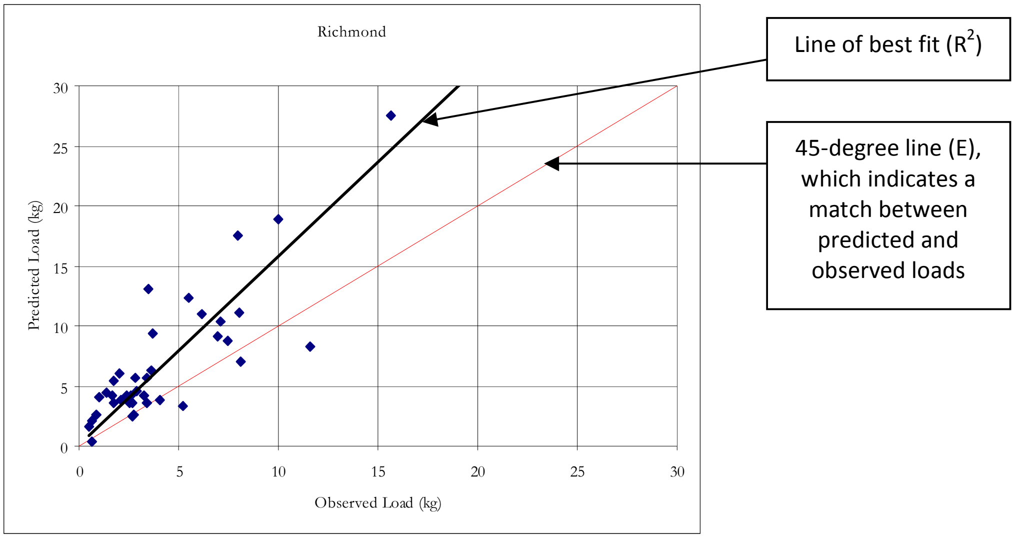


Figure 5-2 Sample results that return a high R^2, but a low coefficient of efficiency (E)

Validation for independent catchments

In a practical sense we want to be able to predict the generated pollutant loads from catchments that have not been intensively monitored, as is the case in most real world management scenarios. The next step in the process was therefore to test the model against catchments independent of those used in calibration.

The method chosen was similar to the cross-calibration exercise above but at a larger scale. Instead of a single event being left out of the analysis, each catchment was removed in turn and a series of calibration parameters was generated using the remaining data sets. By using this method we are verifying the model on a data set (and catchment) that has not been used for model calibration.

Two points must be noted in this analysis:

- To compare different catchments the timestep of the rainfall must be consistent (Term *T* in equation 5.1). The majority of the catchments, in particular the primary data sets, had a minimum of a one-minute timestep and therefore this analysis was undertaken only for those catchments with corresponding 1-minute rainfall. Several datasets were excluded from this phase of model development due to differences in monitoring approach or quality of data collected (Table 5.1)

- Because the model is developed by optimising *b* before the *W* term is calculated (see Figure 5-1), *b* must be consistent for each pollutant (the relationship between *b* and *W* will be further analysed and discussed later in the Chapter). Therefore an optimal *b* was calculated at all catchments (by optimising the line of best fit for each rainfall function/load plot) for each pollutant. By averaging all coefficients of regression, an “optimum” *b* was obtained – that which received the highest average R^2. The optimal R^2 value was then used to determine *W* for the calibration set used

The independent catchment methodology therefore needs a relatively large data set to give meaningful results. In addition to data sets with consistent timesteps and pollutants, a sufficient number of data points are needed. Inconsistent data sets, or data sets with few events were therefore excluded from this analysis (Table 5-1).

Table 5-1 Data sets removed from independent catchment analysis, along with justification for exclusion

Catchment/Pollutant	Reason for exclusion
Monash Roof/TSS, TP,TN	Initial results show a much greater variance in the load – rainfall function relationship, particularly if all events are included (see Chapter 4). The EMC values recorded are also significantly different than the other sites. It is likely that the Roof is a special case for which a specific set of parameters would be applicable.
Burwood East/TP, TN	Less than 15 events in dataset
Blackburn Lake/TSS, TP, TN	Rainfall data only has a timestep of 6 minutes; all other sites have one minute rainfall data. Blackburn Lake data will be considered when 6-minute data is used (Section 0)
Narre Warren TN, TP	As discussed in Chapter 4, the TP and TN loads from Narre Warren are affected by septic tank connections in the catchment.
Lund/TSS	Different monitoring regime with Turbidity used as a surrogate for TSS. As shown in Chapter 4, the rainfall conditions for Lund are also significantly different to those in Eastern Australia
Belgrade/TSS	Different monitoring regime with Turbidity used as a surrogate for TSS.

To gain further insight into the processes involved, the calibration methodology is applied in two ways, one using the total catchment area as the *A* value in Equation 5 -1, and the other using only the impervious catchment area. If the impervious area results are significantly better than the total area results it would indicate that these impervious areas dominate the total loads in the catchments selected (or that the data sets do not contain enough large events which generate pervious area runoff to influence the results) (Boyd, Bufill et al. 1994). It is possible, even likely, that the data sets are not extensive enough to conclusively inform the study about the effects of pervious area. By their nature there are few events monitored which are large enough to generate pervious area runoff, particularly in the largely urban catchments used in the current study.

Sensitivity to model parameters

Given the method of model development with b being determined by optimising the line of best fit, and W representing the slope of the graph, it is probable that the two calibration parameters are not independent. They are also likely to be influenced by the method of data collection – the timestep and rainfall tip size. The goal of this analysis was to look at the interactions between the parameters, and examine if they were unique for a given set of input data.

Many models in general (Ginzburg and Jensen 2004) and stormwater models in particular (Kanso et. al. 2004) have come under criticism due to the complexity of the model outweighing the data available for use. Indeed, some stormwater quality models have been shown to contain parameters that add little to performance (Gaume, Villeneuve et al. 1998), whilst requiring additional data to satisfy the calibration requirements. Any proposed model needs examination to ensure that model structure is appropriate to both the task at hand and the data available, and to determine the differences in model uncertainty of alternative model structures.

Sensitivity to timestep

The model parameters calculated in the preceding sections are dependent on the model timestep. It has been postulated by Duncan (personal communication) that the optimum timestep for prediction will be dependent on the size of the catchment. Initially a timestep of 1-minute was selected (with a rain gauge tip size of 0.2 mm) for the Primary data sets as this was:

- Relatively cost effective and practical to maintain during the monitoring program
- Estimated to provide a high degree of definition about the rainfall intensity

Two sites were available with timesteps and tip size of finer detail than 1 minute and 0.2 mm – Belgrade and Lund. Several sites, and some events within individual sites, were available with only 6-minute rainfall data. In practice in Australia, 6 minute data is widely available for a great number of locations. Therefore, a model that gave reliable results at a 6-minute timestep would be widely applicable in practice.

To test the impact of differing timestep, rainfall was aggregated to reflect a larger timestep at each site. At Belgrade and Lund timesteps of 15 seconds, 1 minute and 6 minute were trialled and the results compared against the correlations obtained for the other sites. At the Primary data sites, and Sandy Ck, the one-minute timestep was compared to a rainfall function derived using a 6, 12 and 30-minute timestep. The relationship between timestep and tip size is also discussed, with Sandy Creek having a tip size of 1mm and Lund having a tip size of 0.035mm.

Independence of Parameters

Given that the proposed model produces consistently high correlations with event loads (Chapter 4), analysis was undertaken to determine the independence of the calibration parameters *b* and *W*. This analysis aimed to examine if the solutions presented by the model were unique solutions or just one of a series of several possible solutions dependent on the calibration parameters. The result also provide an indication of the validity of the model structure itself, as a model that was insensitive to the parameter set may indicate that the structure of the model or number of parameters was not optimal.

Every combination of *b* and *W* values (within a wide range selected by trial and error to ensure that multiple solutions were not occurring) was used to predict an event load. The strength of the relationship of each combination was tested with the Nash-Sutcliffe coefficient of efficiency “E”. Results were plotted on a three-dimensional plot with the *b* and *W* values on the x and y axis, and the Nash-Sutcliffe E on the z axis.

A distinct peak in this case would indicate that there is a unique solution of the model parameters and that the parameters are independent (Gaume, Villeneuve et al. 1998). A plateau or a series of peaks would indicate a variety of solutions or parameters that are co-correlated (and thus dependent on each other).

5.4 Results

Examination of rainfall function/event load plots

The visual examination of the rainfall function versus load plots was aimed at two specific factors – any indication of curvature in the plot and the significance of the y- intercept (Figure 3-2 and Appendix A). Curvature would indicate the model formulation would need to be reconsidered, however despite scatter in the points, there is little evidence of curvature in the relationships.

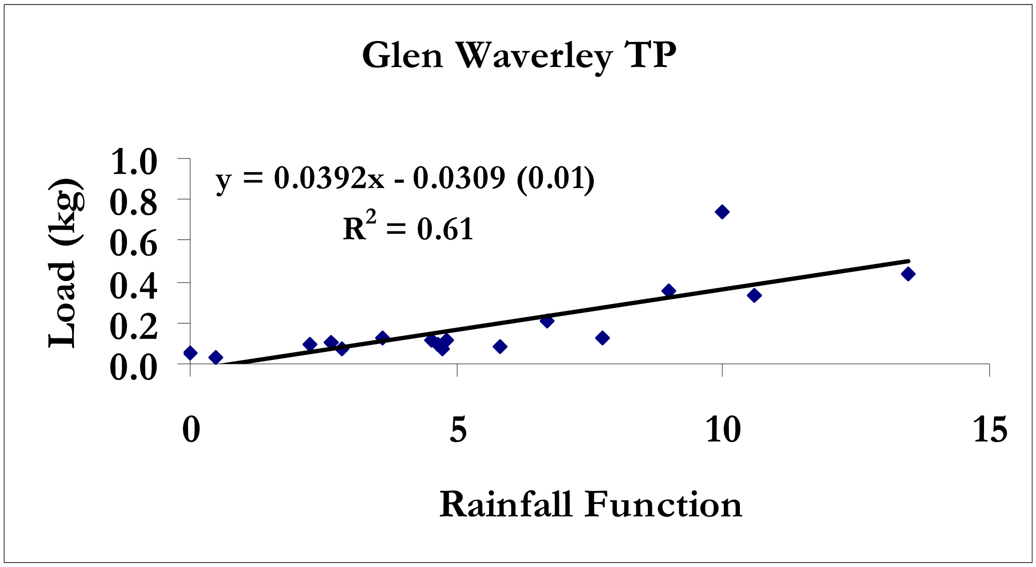


Figure 5-3 Example plot of rainfall function versus pollutant load (with significance of the y-intercept in brackets)

Associated with this, from visual inspection, the Y intercept values tend to be very small (see Appendix A). The significance of the Y intercept at the 95% level was examined for all cases. The Y intercept is significantly different from zero at the 95% level in around half the cases, with approximately half of these having negative y-intercepts.

Although this is not conclusive it is proposed that the straight-line regression relationship can be further simplified to a straight line through zero. The practical implication of this is that when the rainfall function input (and hence the storm volume) is zero, the pollutant load is zero. If the model results indicate that further explanation of pollutant behaviour is needed in either the description of event loads or within event behaviour further examination can be undertaken.

In doing so, we are actively seeking greater simplicity. We do not claim that the extra detail (in this case the Y intercept) has absolutely no effect. We do claim, however, that every extra detail and every extra parameter to be fitted must be paid for with more data to achieve a satisfactory fit, and with more noise introduced along with the additional data (Ginzburg and Jensen 2004). Our approach is to start with the simplest possible formulation compatible with our data. Further detail can be added if necessary as future data demands and permits.

Validation for individual catchments

Because of the significance of the results when the initial plots of rainfall function versus loads were considered the assumption of a zero y-intercept has therefore been used for calibration analysis. When the cross calibration exercise is carried out, total loads at each site are predicted well (Table 5-2). When predicting long term loads there is a maximum percentage error of 18% (TSS at Belgrade), the average error across all sites being 9%, 5% and 5% for TSS, TP and TN respectively.

When the prediction of individual events is considered the method achieves *E* values typically above 0.5 and usually above 0.7 for all pollutants. Exceptions to this are TSS (derived from turbidity) in the case of Belgrade, and TP for the Monash Roof and Narre Warren. TN correlations for the latter two sites are also low. As has been discussed the TP and TN results for Narre Warren are influenced by the significant loads derived from the septic systems of the catchment. The very low E value for TP at Narre Warren (remembering that a negative value indicates that the mean load is a better predictor than the function tested) is due to one event - 35 on the 2/2/05 which has been discussed in Chapter 4 - having a significant influence on the results (removing event 35 causes an increase in E to 0.2 and R^2 to 0.87).

Table 5-2 Results from each catchment for the cross calibration exercise (TP and TN at Eley Road have been left out due to there being only 10 events with one-minute timestep data)

Catchment	Pollutant	Observed/Predicted Total Load	R^2 of Observed vs Predicted event loads	E of Observed vs Predicted event loads
Primary Data Sets				
Mt. Waverley	TSS TP TN	0.99 1.00 1.01	0.84 0.71 0.88	0.82 0.65 0.88
Glen Waverley	TSS TP TN	0.89 0.93 0.96	0.61 0.47 0.67	0.58 0.46 0.67
Narre Warren	TSS TP TN	1.08 1.03 0.95	0.83 0.79 0.85	0.67 -0.04 0.56
Monash roof, Clayton	TSS TP TN	0.96 0.95 1.12	0.86 0.57 0.62	0.82 0.45 0.51
Burwood East	TSS	1.07	0.71	0.69
Richmond	TSS TP TN	1.13 1.11 1.06	0.76 0.70 0.79	0.74 0.68 0.79
Doncaster	TSS	0.92	0.94	0.92
Secondary Data Sets				
Blackburn Lake	TSS TP TN	0.97 1.09 0.96	0.90 0.71 0.83	0.90 0.70 0.82
Sandy Creek	TSS TP TN	1.16 1.05 1.00	0.80 0.93 0.98	0.79 0.92 0.98
Lund	TSS	0.98	0.71	0.71
Belgrade	TSS	1.18	0.59	0.33

Plots of the predicted versus observed loads (Figure 5-3; see Appendix B for TP, TN and secondary data sets) add additional information, especially the influence of large events on the overall correlations. As with the rainfall runoff relationships, three of the TSS data sets are significantly influenced by single large events. The TSS result for Doncaster is an example of this effect, with the February 2005 event already discussed dominating results (Figure 5-3). Where a wide range of events were available, however, the model provides a good representation of event loads over the range. It is also important to note that the large events carry a significant proportion of the total load, and therefore are of primary importance in situations when total loads are of significance to the receiving environment. As an example, the largest event at Doncaster carries 30% of the total TSS load for the 54 events sampled.

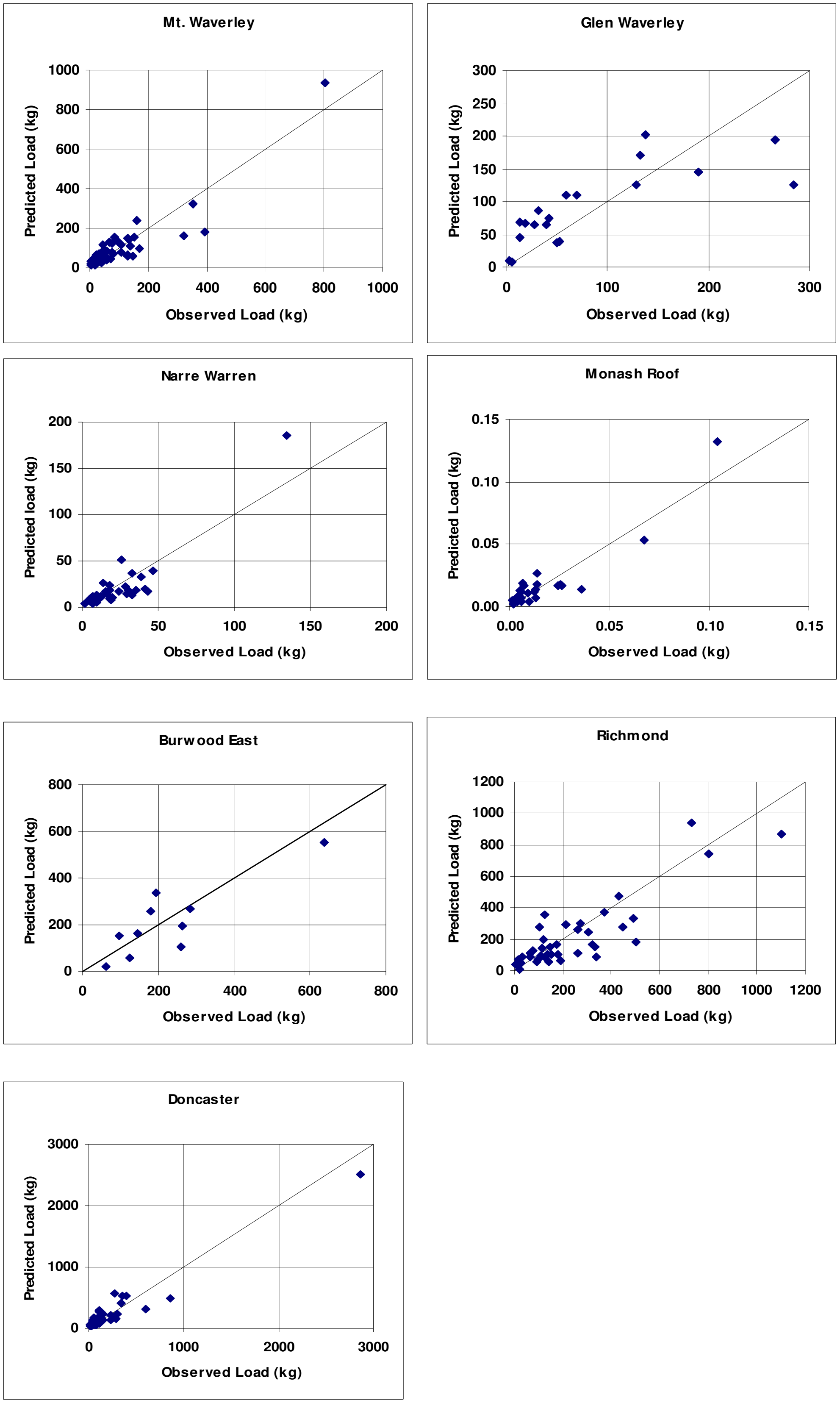


Figure 5-4 Predicted versus Observed loads (TSS)

Validation for independent catchments

As discussed in the introduction to this chapter, many models have adequately predicted total loads when calibrated against a particular catchment. A more thorough and important test of the model is its ability to predict loads for catchments not used in the calibration, as this is likely to be the situation faced in most real world situations.

The optimum *b* for the rainfall function equation shows distinctly different values for each of the pollutants. To optimise b for each pollutant the correlation between load and rainfall function was trialled using a wide range of b values which were then plotted (Figures 5 -5, 5-6, 5-7). The optimum figure was taken as the average of the best *b* values for all catchments (Table 5-3).

The optimum b for each pollutant is 1.8 for TSS, 1.4 for TP and 1.2 for TN for a one minute timestep.

Table 5-3 Optimum “b” for TSS, TP and TN

Catchment	Optimum “b” in Rainfall Function equation (one minute timesteps)		
	TSS	TP	TN
Mt. Waverley	1.6	1.2	1.3
Glen Waverley	1.3	1.2	1.2
Narre Warren	1.8	1.7	1.4
Monash roof, Clayton *	2.6	1.1	1.1
Burwood East	1.4		
Richmond	2.3	2.0	1.6
Doncaster	1.9		

* events 18 and 23 removed

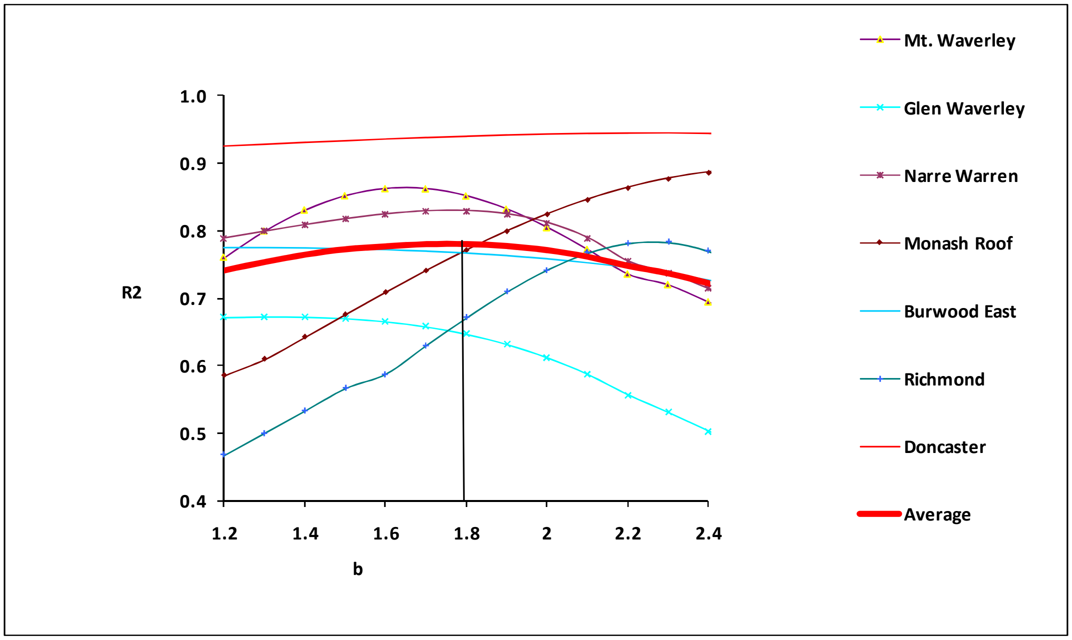


Figure 5-5 Value of b vs. Correlation (R^2) - (TSS), the vertical black line represents the optimum b figure

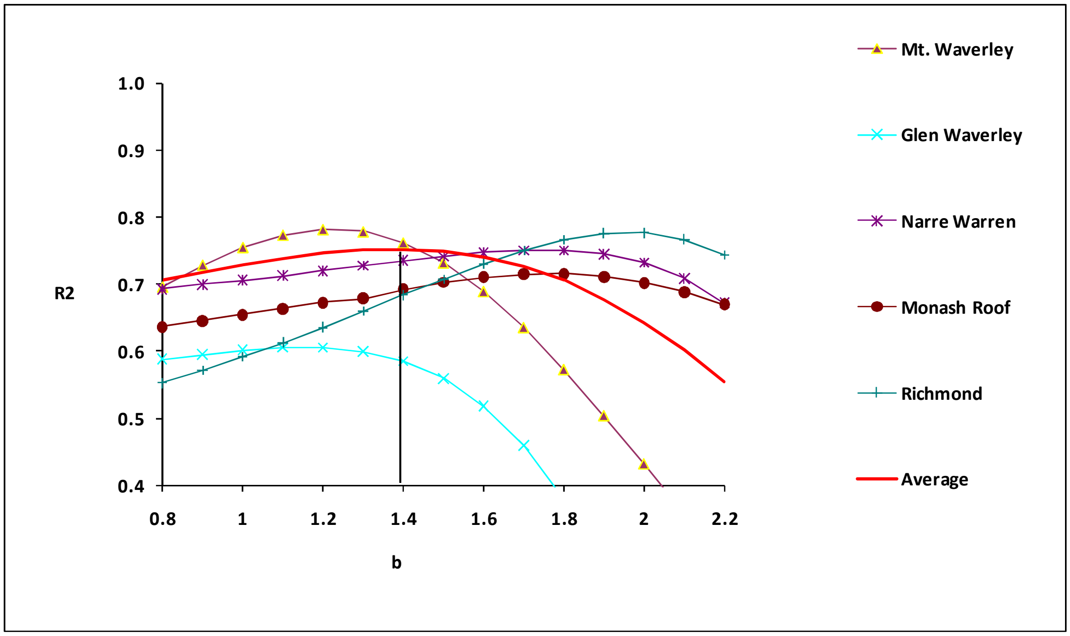


Figure 5-6 Value of b vs. Correlation (R^2) - (TP), the vertical black line represents the optimum b figure.

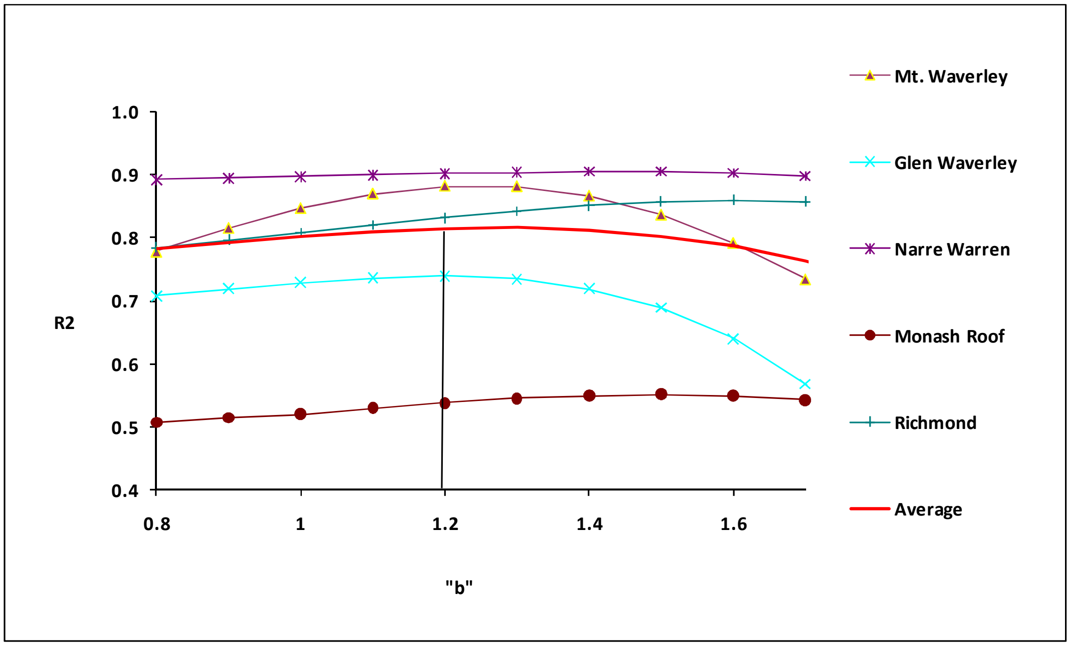


Figure 5-7 Value of b vs. Correlation (R^2) - (TN), the vertical black line represents the optimum b figure.

Given the optimum *b* values, the corresponding *W* value could be calculated for each catchment (Table 5-3). During the individual catchment analysis each catchments had its own unique combination of *W*, *T* (timestep) and *b*. When comparing catchments, however, the b value was kept at the optimum for each pollutant, and only catchments with consistent timesteps(one-minute) were used. The analysis was undertaken using both the total catchment area and the impervious area of the catchment (

Table 5-4).

	Catchment	Corresponding "W" when A =total area	Corresponding "W" when A =impervious area
TSS	Mt. Waverley	0.92	1.19
	Richmond	1.21	1.64
	Roof*	0.48	0.48
	Glen Waverley	0.43	1.08
	Narre Warren	0.42	2.08
	Doncaster	0.91	1.78
	Sandy Creek	0.70	1.57
	Burwood East	0.13	0.29
TP	Mt. Waverley	0.0013	0.0016
	Richmond	0.0032	0.0043
	Roof*	0.0004	0.0004
	Glen Waverley	0.0009	0.0023
	Narre Warren*	0.0024	0.0121
	Sandy Creek	0.0021	0.0047
TN	Mt. Waverley	0.0090	0.0115
	Richmond	0.0123	0.0166
	Roof*	0.0036	0.0036
	Glen Waverley	0.0070	0.0176
	Narre*	0.0103	0.0516
	Sandy	0.0138	0.0306

When the cross catchment calibration exercise is conducted the predictions for total loads are generally good (Table 5.5), with the maximum difference for long term loads being 52% for TSS at Mt. Waverley. The average differences are 19%, 16% and 30% for TSS, TP and TN respectively.

When individual events are considered, the coefficient of efficiency is generally above 0.6, however TP and TN at Glen Waverley and TN at Richmond returned negative values.

	Catchment	Corresponding "W" when A =total area	Corresponding "W" when A =impervious area
TSS	Mt. Waverley	0.92	1.19
	Richmond	1.21	1.64
	Roof*	0.48	0.48
	Glen Waverley	0.43	1.08
	Narre Warren	0.42	2.08

	Doncaster	0.91	1.78
	Sandy Creek	0.70	1.57
	Burwood East	0.13	0.29
TP	Mt. Waverley	0.0013	0.0016
	Richmond	0.0032	0.0043
	Roof*	0.0004	0.0004
	Glen Waverley	0.0009	0.0023
	Narre Warren*	0.0024	0.0121
	Sandy Creek	0.0021	0.0047
TN	Mt. Waverley	0.0090	0.0115
	Richmond	0.0123	0.0166
	Roof*	0.0036	0.0036
	Glen Waverley	0.0070	0.0176
	Narre*	0.0103	0.0516
	Sandy	0.0138	0.0306

Table 5-4 Optimum values of W for all catchments (consistent b and T) using both total and impervious area

Table 5-5 Independent catchment calibration exercise using total catchment area

Catchment	Pollutant	Observed/ Predicted Total Load	R2 of Observed vs Predicted events	Coefficient of Efficiency of Observed vs Predicted events
Sandy Creek	TSS	0.88	0.86	0.75
	TP	1.56	0.96	0.80
	TN	1.45	0.96	0.87
Mt. Waverley	TSS	0.64	0.65	-0.12
	TP	0.61	0.76	-0.49
	TN	0.48	0.88	-1.26
Glen Waverley	TSS	1.33	0.65	0.48
	TP	0.93	0.56	0.56
	TN	0.76	0.74	0.56
Narre Warren	TSS	1.66	0.86	0.67
Richmond	TSS	0.95	0.95	0.94
	TP	0.48	0.66	-2.52
	TN	0.30	0.82	-11.5
Doncaster	TSS	1.17	0.89	0.88

Table 5-6 Independent catchment calibration exercise using impervious catchment area

Catchment	Pollutant	Observed/Predicted Total Load	R^2 of Observed vs Predicted events	Coefficient of Efficiency of Observed vs Predicted events
Sandy Creek	TSS	1.05	0.89	0.88
	TP	1.09	0.96	0.96
	TN	1.05	0.96	0.94
Mt. Waverley	TSS	1.52	0.86	0.75
	TP	0.87	0.76	0.58
	TN	0.73	0.88	0.66
Glen Waverley	TSS	1.06	0.65	0.61
	TP	0.59	0.56	-0.09
	TN	0.51	0.74	-1.10
Narre Warren	TSS	0.94	0.86	0.69
Richmond	TSS	1.08	0.65	0.64
	TP	0.99	0.66	0.61
	TN	0.61	0.82	-0.25
Doncaster	TSS	0.99	0.95	0.94

When only the catchment impervious area is considered the results are similar, but slightly inferior to those using total area (Table 5-6). Average differences of the total load for all events were 28%, 38% and 48% for TSS, TP and TN respectively.

This result is somewhat expected, as the rainfall function vs load analysis in Chapter 4 shows no apparent curvature which would indicate a significantly different pollutant generation mechanism. The use of impervious area alone does not appear to add value to the model, at least for the data available. There is a possibility that further data, particularly for large events that included pervious area runoff, would add further insights into the process. Given the similar results between total and impervious area the use of impervious area does not appear warranted. The use of impervious area adds an extra source of error into the measurement as impervious fraction needs to be estimated in addition to total area. Given this, it is proposed that further model development be undertaken using total area.

Sensitivity to model parameters – model timestep

One of the key points of investigation when developing the monitoring program was to investigate the impact of timestep on model performance. Chapter 3 describes previous attempts to describe pollutant generation using rainfall intensity using long timesteps, or event averages of the rainfall intensity. The relationship between the timestep and the rainfall tip size impacts on model results.

The results do indeed show that generally the strength of the correlations decreases as the timestep increases (Table 5-8). Indeed, as expected, the one-minute data generally gives the best correlations, although in all cases the difference is relatively small.

The Lund and Belgrade catchments are particularly interesting. In both cases the small timestep gives inferior results (Table 5-7) to a one or six minute timestep. This may be explained through the timestep – rain gauge tip relationship. By using the rainfall function, the sum of the rainfall intensity to a power at each timestep, we are attempting to represent the input of kinetic energy through rainfall impact. Therefore a timestep that averaged the rainfall would lose some of that information. Conversely, when a very small timestep is used (such as the 15 second timestep at Belgrade and Lund) there is little chance that more than one rainfall tip will be recorded in a timestep. As the power function has most influence when multiple tips are recorded in a timestep this would again reduce the usefulness of the model.

Sandy Ck is insensitive to the parameter b (Figure 5.8). The Sandy Ck dataset has the largest rainfall gauge tip size (one millimetre) which limits the definition of rainfall intensities that can be measured. For many events, and timesteps within an event, the measurement becomes binary – either on or off. When raised to a power, such as in Equation 5.1, the event tends to become purely a sum of rainfall tips rather than a reflection of varying rainfall intensity. Similarly the Lund and Belgrade sites have very small tip sizes over very short timesteps, and a similar effect occurs where the relationship between tip size and timestep affects the ability of the model to provide definition. Optimum timestep is a function of tip size, and to a lesser extent the range of rainfall intensities observed.

Table 5-7 Effect of different timesteps at the Lund and Belgrade sites

	R^2 of $W\Sigma I^b$ and TSS Load at Timestep			
Catchment	15 seconds	1 minute	6 minutes	30 minutes
Lund	0.63	0.76	0.77	0.76
Belgrade	0.42	0.72	0.70	0.62

Table 5-8 Effect of different timesteps at Australian sites

	R^2 of $W\Sigma I^b$ and TSS Load at various timesteps			
Catchment	1 minute	6 minutes	12 minutes	30 minutes
Blackburn Lake		0.94	0.89	0.88
Sandy Ck	0.87	0.89	0.86	0.84
Mt. Waverley	0.91	0.90	0.82	0.80
Glen Waverley	0.83	0.83	0.68	0.72
Narre Warren	0.91	0.91	0.93	0.91
Burwood East	0.93	0.88	0.83	0.83
Roof (Events 18 & 23 removed)	0.89	0.87	0.73	0.69
Richmond	0.90	0.86	0.84	0.81
Doncaster	0.95	0.96	0.93	0.47

The practical result of these figures is that, at least for the majority of conditions covered by this monitoring program, six minute rainfall data will provide comparable results to a one minute timestep, given the 0.2 mm rainfall gauge tip size. Given that this combination is a common one in Australia, a set of calibration parameters was calculated for the primary data sets using the same methodology as for the independent catchments analysis, but with a six minute timestep.

Table 5-9 Calibration parameters based on a 6-minute timestep (0.2mm tip size)

Pollutant	b (6-minute timestep)	µW	σW
TSS	1.5	0.37	0.25
TP	1.3	0.0012	0.00084
TN	1.0	0.0063	0.0037

Independence of parameters

It was seen in figures 5-5, 5-6 and 5-7 that the model has limited sensitivity to the power parameter *b*. For instance any value of *b* between 2.2 and 1.3 will give an average R^2 of over 0.75 (the corresponding range for TP is 1.6 to 1.2 and 1.7 to 0.6 for TN). The parameter *b* does however clearly vary between the pollutants. Given that the model is performing at least as well as others reported in the literature, further investigation was undertaken to examine both the sensitivity of the model to its calibration parameters and the interdependence of those parameters.

Initially the three dimensional plotting exercise shows the strength of the relationship (E for various values of *b* and *W*. All plots showed a distinct peak or 'ridge' (**Figure 5-8**) which is a function of the relative insensitivity to the power parameter *b*. As expected the shape of the peak reflects the plots shown in figure 5-5.5-6 and 5-7 of R^2 versus *b*. Although most of the plots show this 'ridge', there are no cases where multiple peaks or ridges appear which would indicate a different *b*/*W* combination.

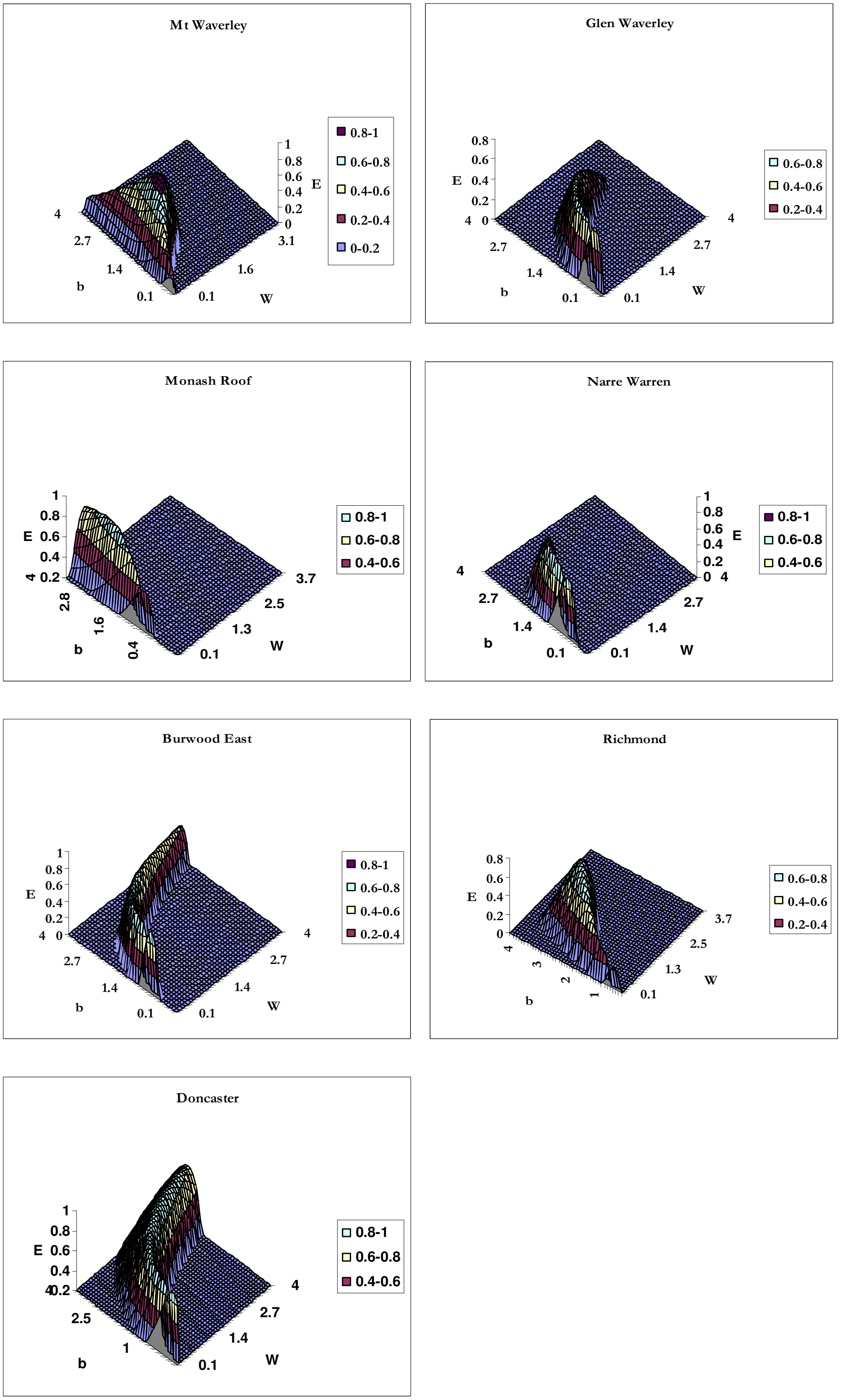


Figure 5-8 Sensitivity of model parameters, Primary catchments (TSS)

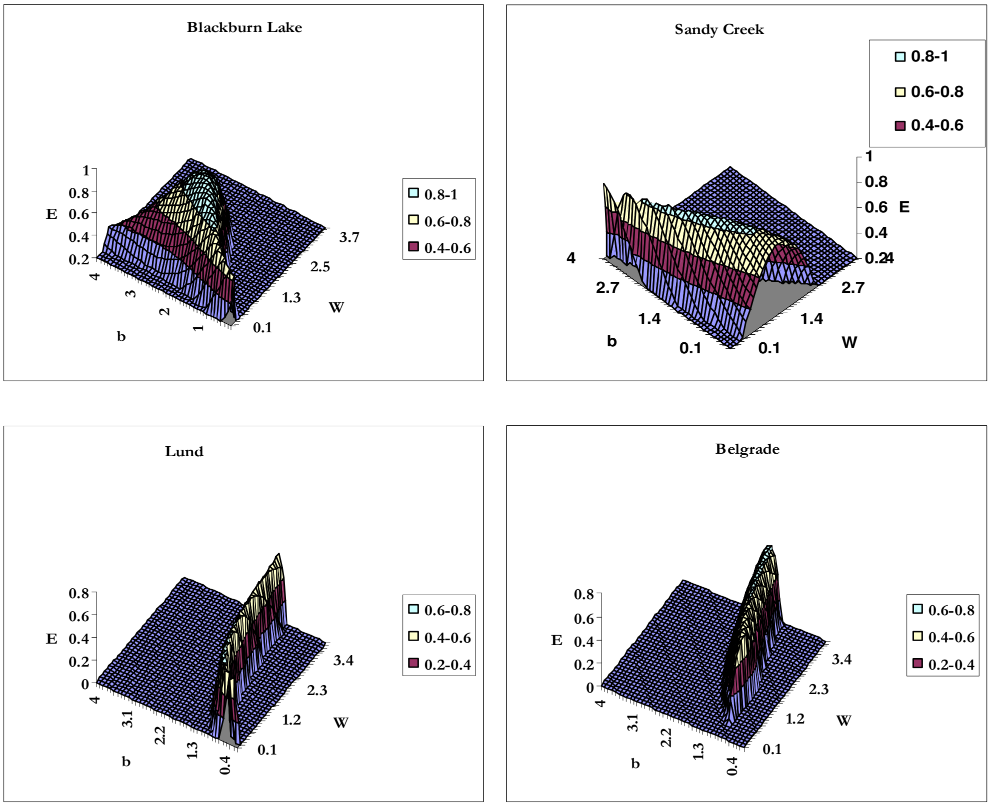
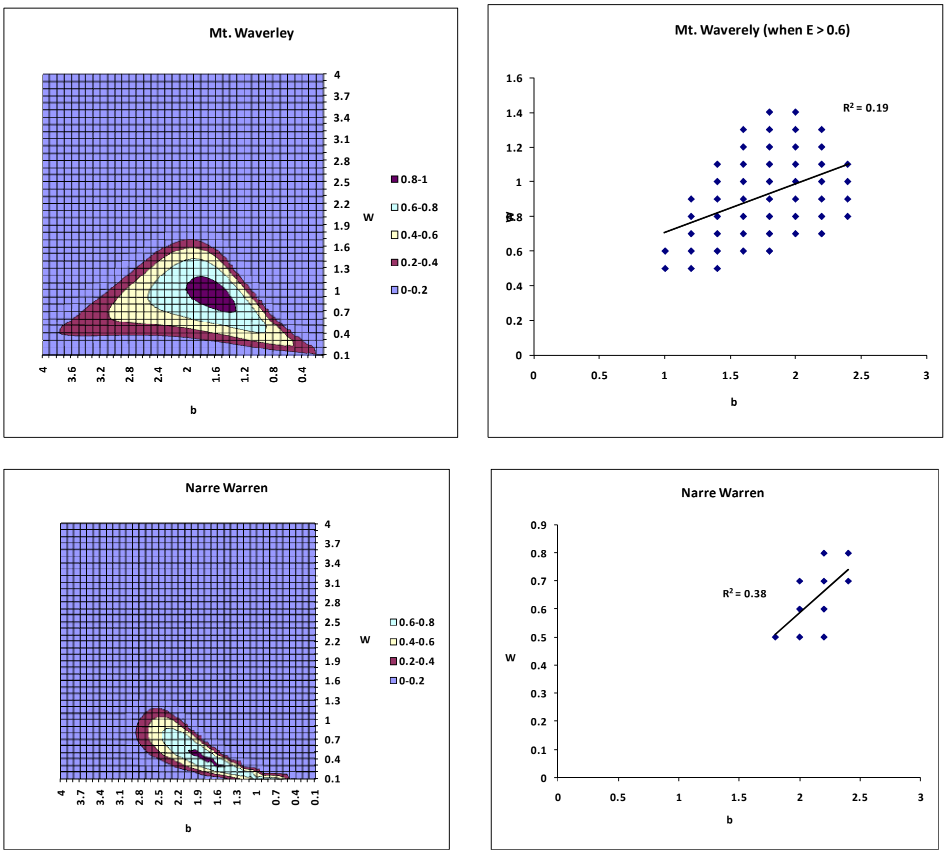


Figure 5-9 **Sensitivity of model parameters, Secondary catchments (TSS)**

If the calibration coefficients *b* and *W* were completely independent we would expect each of the datasets to present distinct peaks as opposed to ridges. Indeed, several of the catchments do demonstrate this. To attempt to quantify that relationship, each of the plots was represented using a two dimensional view, with the strength of the E value represented by different colours. Each combination of *b* and *W* that achieved an *E* value of greater than 0.6 was plotted on a two dimensional plot (Figure 5-9). This is in a sense “looking down from above” on the three dimensional plots. Most datasets show a distinct relationship between the coefficients (Figure 5.10 and Appendix C), indicating that the coefficients are interrelated given the data available.

In general the sites with a better spread of events (ie. a range of small, medium and large storm events) tend to be more sensitive to *b* and therefore show a more distinct solution. In this sense the sensitivity analysis does not tell us more than we already know – that sites dominated by a few large events or small storm events with timesteps/tip sizes that limit definition – are insensitive to b and therefore show more interdependence between parameters. Of course the data sets that are dominated by one large event are themselves distorted, or inadequate samples to represent fully their underlying populations.


Mt. Waverley

Mt. Waverely (when E > 0.6)

Narre Warren

Narre Warren

Figure 5-10 Mt. Waverley and Narre Warren TSS sensitivity to model parameters

5.5 Discussion

Model Calibration

The method itself – using rainfall intensity at short timesteps - consistently improves predictions when compared to total volume, or a function of the runoff across all the various climatic zones covered. The results support the hypothesis that rainfall impact energy is an important factor in the generation of at least some pollutants. The power term in equation 5.1 is higher for TSS than TP, which is in turn is higher than for TN, reflecting the diminishing role of rainfall impact energy as the dissolved proportion of a pollutant increases.

The results are similar to those published comparisons of modelled versus predicted loads. Tsihrintzis and Hamid (1998) calibrated SWMM to Southern Florida catchments and then compared modelled results to the results collected in the NURP program for the same region. TSS EMC predictions were on average 2.5 times higher than measured for both residential and commercial land use. TKN results were about a third of the measured results for residential and two thirds of measured results for commercial land use. All results were considered acceptable if they were in the same order of magnitude as measured results. Driver and Troutman (1989) used regression models in the US (different regression models for 3 different climatic zones and 11 pollutants) to predict long term loads and report prediction errors of 56 to 334% for EMCs. Interestingly, TSS gave the worst results with TN the best. Total rainfall and catchment area were the primary inputs. This result is consistent with the results of the current study, which suggest that a *b* value of 1.0 is the optimum for TN (indicating total rainfall) while 1.5 is the optimum for TSS (Table 5-9).

Indeed the power term optimises at one for TN in the current study, indicating that total rainfall is the most important factor in the case of this, largely dissolved, pollutant. A power of one is still important however, as the proposed model defaults to a linear model under this case.

The Melbourne sites from the current study, with the exception of the Roof site, also produce consistent calibration co-efficients which means they can potentially be used across nearby catchments. These results confirm the hypothesis of Duncan (1995) and early work done (Chiew, Duncan et al. 1997; Smith 1997; Vaze and Chiew 2003) which suggested that a simple modelling approach can achieve as good or better results for pollutant loads than more complex models. Further work is required to establish parameter sets for other climatic zones.

The Roof catchment appears to perform distinctly differently to most others, and although the original correlations between rainfall function and loads are high, more investigation is needed to develop an appropriate parameter set for these particular situations.

Applying the model using total area gives slightly better results than the impervious area method which suggests that, at least for the data sets used, pervious area has not impacted on the generated loads. The results are comparable, however, so it is possible that further studies may provide insights into this topic, particularly if the study catchments with lower fraction impervious are considered. At the present time, given the extra uncertainties involved with calculating impervious area, the total area method is preferable.

Sensitivity of model parameters

The data set used provides significant insights into the model, and suggests reasons why previous approaches have achieved limited success. There is a relationship between time step, tip size, and rainfall intensity that is important when making predictions. When the tip size is large compared to the timestep (as was the case with Sandy Creek), information is lost as there are too few tips in a timestep to provide definition. This impacts upon the model by making it insensitive to the value of *b*. Similarly, if the timestep is too short there are not enough tips in each timestep to provide definition. These factors are in conflict with the need for short timesteps and small tip size to provide a range of rainfall intensity information so that the power function can be fitted. When timesteps are too large, or intensity data is averaged, the model results similarly suffer.

For a model of this nature to succeed, an optimum balance of tip size and timestep is needed. The one-minute timestep, 0.2 mm tip provides good results in the data presented, and the 6-minute, 0.2mm tip combination also produces encouraging results.

The combination of timestep and tip size also has some effect on the *b* parameter. Importantly in all cases however, the optimum *b* for each pollutant is distinct and well defined. This is consistent with findings about models with similar structure such as SWMM (Sriananthakumar and Codner 1992) which use a runoff based term raised to a power. This result provides more weight to the hypothesis that energy from rainfall drops is a critical factor in pollutant mobilisation as the largely particulate pollutants, such as TSS, have a greater *b* value than those that have a significant dissolved fraction.

When the parameters are plotted against one another for all sites there is also a correlation between the *b* and the *W* parameters. It is unclear if this is a fundamental aspect of the model or a result of the data sets tested. Further work on this topic using subsets of the collected data is currently being undertaken by other researchers (Dotto 2009). Initial efforts have focussed on the coupling of a derivation of the proposed model with a rainfall runoff model and attempting to predict time series of concentrations. Although the runoff model can achieve predictions with high correlations to observed pollutant load data, the pollutant generation model is hampered by the co-dependence of parameters b and W. For now the parameters must be viewed as a set. Further data collection in other studies will help to explore this issue.

5.6 Conclusions

The results presented in this Chapter confirm the hypothesis in Chapter 3 that by representing the physical process of rainfall intensity striking urban surfaces, the pollutant loadings can be well predicted. Short timestep rainfall intensity appears to be an important contributing factor to pollutant generation. The $\sum$ term in the model acknowledges that the total flow or rainfall gives reasonable results, however the intensity terms increase the accuracy of the predictions. The predictions of event loads for TSS for independent catchments are reasonably good within the uniform climatic conditions represented by the monitoring program. Predictions of TP and TN are also good, but more variable. Most importantly the model offers hope that, within a given climatic zone, pollutant loads can be predicted well across unmonitored catchments.

When the structure of the model is interrogated, it appears that the relationship between timestep and tip size is critically important and is probably a contributing factor to the mixed success of this type of model in the past. The results from Queensland (Sandy Creek) and Europe (Lund and Belgrade) are also encouraging in terms of the simple model, however further data is needed to develop an understanding of the model parameters under different climatic conditions such as these. A set of parameters for Melbourne is presented that will allow comparison with future studies.

6 WITHIN EVENT BEHAVIOUR

6.1 Introduction

Chapter Five demonstrated that the rainfall function can predict event loads with accuracy as good as, or possibly better than published results from previously described complex models (Section 5-5). It also shows that rainfall intensity, at least for Melbourne climatic conditions, is a better predictor of loads than is runoff rate, total rainfall or total runoff. It does not, however, describe how the pollutants will be generated within an event, or how the behaviour of pollutants varies between parameters.

The analysis has, however, added insight into the primary drivers behind pollutant washoff, which for medium sized urban catchments within the climatic zone tested, appears to be driven by rainfall intensity. This chapter will examine if those insights can be transferred into practical information at a short timescale, to describe temporal variations within an event.

If rainfall intensity is the primary driver for event loads it should also be the driver for changes in incremental loads during an event. Many of the complex models described in previous chapters reply on exponential decay concepts of washoff, and sometimes buildup. By definition these models make assumptions about the pollutant behaviour during an event. Many authors (Alley et al. 1980; Huber and Dickinson 1988; Sriananthakumar and Codner 1992; Vaze and Chiew 2003) have recognised that these assumptions impose limitations on the applicability of the models.

Chapter 6 therefore has two primary aims:

- To describe the pollutant behaviour during the events recorded in the monitoring programs, and therefore to describe if behaviour such as *first flush* and *depletion of built up pollutants* is occurring within the monitored catchments.

- To investigate the proposed rainfall function's ability to simulate the observed behaviour at short timesteps.

6.2 Background

The important aspects of pollutant behaviour within individual events are the presence or absence of trends that lead to greater understanding of physical processes. Many authors have also described a "first flush" effect, where disproportionate loads of pollutants are carried by the early flow in an event. This concept is appealing to practitioners as it offers the possibility of treating a small amount of flow but a major proportion of pollutants. This appeal is played out in many stormwater treatment designs that have a specific goal to treat the first flush. Real assessments of the usefulness of the concept to stormwater treatment aims are rarely reported.

Also associated with variations of pollutant loadings within events is the concept of depletion of pollutants with time or storm volume over an event. Chapter 2 discusses the strengths and weaknesses of this approach. A traditional understanding of the process would see runoff gradually depleting the pollutants available for washoff, and the resulting stormwater becoming "cleaner" over time. As discussed in chapter two, however, experimental and theoretical studies have questioned this approach, and suggested that in the vast majority of rainfall events the pollutants on the surface far exceed the possible washoff loads (Malmquist 1978; Vaze and Chiew 2003b).

Much has been written about the concept of "first flush" in the scientific literature, including much debate about the definition (Bertrand-Krajewski et al. 1998; Deletic 1998). Generally first flush refers to a situation where the pollutant cumulative load exceeds the cumulative flow early in an event. First flush has both been claimed to occur (Spangberg and Niemczynowicz 1993; Deletic 1998) and not to occur (Deletic 1998) for the same data sets, which indicates the lack of an agreed definition or process description. Many definitions have been used (Bertrand-Krajewski et al 1998; Deletic 1998), the most meaningful relating the phenomenon to possible treatment methods (Kang et. al. 2006).

Depending on the definition, first flush has been demonstrated to occur on small catchments (Yaziz et al. 1989), and in some cases related to the antecedent dry period (Mason et al. 1999). Other small catchments described have not shown a distinct first flush behaviour, at least for TSS (Deletic 1998). As catchment size (and time of concentration) increases first flush becomes less clear and less relevant (Vorreiter and Hickey 1994; Chiew et al. 1997; Cristina and Sansalone 2003). It is, by its nature, a feature of small catchments. In large catchments differing attenuation and lag times of the various sub-catchments negate any distinct effect (Duncan 1995).

Within event behaviour is usually demonstrated through the use of pollutographs (Deletic 1998) – plots of the predicted flow rate or concentration of pollutants compared with observed data. Pollutographs provide a visual representation of the pollutant flux, but they do not, on their own, provide quantifiable data about the representation. Some authors have used mathematical methods to measure the time between observed and predicted peaks (Kang et. al 2006) usually in terms of dynamic calibration of complex models.

Some authors have questioned the need for prediction of short term pollutant behaviour (Huber 1992) and, unless there are strong effects such as first flush that can result in changes to treatment measure design, it is difficult to envisage a situation where that level of prediction is needed by stormwater managers. The main benefit therefore is likely to be added confidence that the model is representing the central processes involved (Kang et. al. 2008), opposed to acting solely as a well calibrated “black box”.

6.3 Method

The chapter analysis is split into two main themes;

- Description of any trends of within event behaviour observed during the monitoring program.

- Assessment of the model’s performance in replicating that behaviour.

Cumulative pollution load versus cumulative volume

The first step taken in examining pollutant behaviour was to examine the pollutant loads relative to flow that occurred in the catchment. This analysis would provide insights into behaviours such as first flush and depletion.

This was initially performed using graphical methods (Figure 6.1) where the cumulative flow and cumulative pollutant load were calculated and then normalised (divided by the total volume and total load respectively).

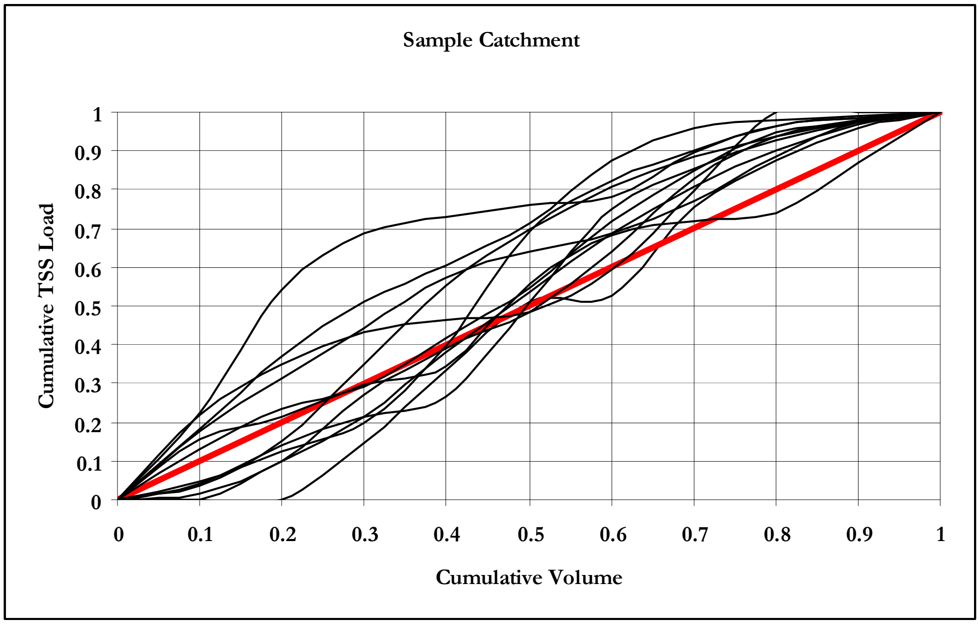


Figure 6-1 Sample cumulative volume/pollutant load plot (where the thick red diagonal line reflects the cumulative volume over an event and the black lines are the cumulative loads for each individual event)

The cumulative load was then plotted against the cumulative volume (sample shown in Figure 6.1). The thick red diagonal line reflects the cumulative volume over an event; if the pollution load was evenly distributed with volume the fine black lines would mimic this line. A slope of greater than 45% (black line above the red line) at the beginning of an event indicates that there is relatively more pollution being washed off – a “first flush”.

To quantify the effect an indicator called the mass first flush ratio (MFF_{20} – the 20 indicating that 20% of the flow has passed) is used (Ma et al 2002). This is the ratio of the percentage of load carried divided by the percentage of flow (20% in this case). Therefore if 40% of the total pollutant mass is carried by the first 20% of flow the MFF_{20} would be 2.0. As discussed in the background to this chapter a “first flush” is largely of interest from a practitioner point of view if it can be used in the design of pollution mitigation measures. A high MMF value would leave open the possibility of designing treatment facilities to treat only the first portion of flow.

The analysis should determine if there is a pronounced first flush in any or all of the monitored catchments and give some indication of the strength of the phenomenon. It will also aim to determine any differences in the effect across the catchments.

Depletion of pollutants with time and volume

Many physical models assume that there is a reduction in pollutant washoff with time due to less pollution being available on the surface later in a storm (Huber and Dickinson 1988). The practical explanation is that, as a storm progresses, there is less material (pollutant) available on the surface and therefore the amount washed off decreases. The depletion term is usually expressed as being relative to some other factor such as storm volume or duration.

As we have shown that rainfall intensity is a strong explanatory variable in pollutant loads (Section 5-4) therefore we want to test if there are other factors contributing to the pollutant behaviour within an event. In our case the question becomes: “if the rainfall intensity is constant during an event, does the washoff rate of pollutants or the concentration alter, and if so how?”

By answering the question above we can assess whether there is a need for a depletion term in the model (Equation 6-1).

$$Load = W \times A \times T \times \sum_{i=1}^{n} I^{b}$$ **Equation 6-1**

where *Load* is the load of pollutant produced by the whole catchment over the period of interest (units of mass), *W* is a calibration coefficient dependent on pollutant and location, *A* is assumed to be the total catchment area (units of length^2), *T* is the duration of the timestep (units of time), *I* is the rainfall intensity in the timestep (units of depth/time), *b* is a constant dependant on the pollutant, and *n* is the number of time steps over the period of interest.

For a given event (with constant area A and timestep T) the *W* term becomes a ratio between the load and rainfall intensity (and therefore predicted load). The variation in ΣI^b over an event therefore becomes a measure of how the washoff load varies with intensity. If this ratio changes consistently with duration or storm volume it indicates other variables may need to be considered to explain the behaviour.

To perform this analysis every point in time for which a water quality sample was taken an optimal *W* (ratio between rainfall intensity and observed load) value was calculated, by dividing the observed load for the timestep by the rainfall intensity at that timestep. The cumulative storm volume and time elapsed, for the same point in time in the event, was recorded.

For each of the points recorded the optimal *W* was then plotted against the cumulative storm volume and event duration. If pollutants were being depleted, the gap between the measured load and the rainfall function would increase, leading to a decrease in the optimal *W* value over time. The slope of trendlines on these graphs was assessed for its significance (at a 0.05 level).

As an example at the Mt. Waverley site there were over 700 TSS samples used for analysis over the monitoring program. Using the catchment specific calibration of *b*, the rainfall function was compared to each of these values and an optimal *W* value determined. The optimal *W* values were then plotted against both the time elapsed and the cumulative runoff volume.

It is acknowledged that this approach does not explicitly test for depletion, instead it examines the data for a consistent pattern of under or over prediction that would justify the inclusion of more variables in a predictive model.

Comparison of Predicted and Observed Pollutant Behaviour - Calculation of catchment lag times

The obvious limitation with using Equation 6-1 in the prediction of pollutant flux is the lack of method of predicting time lag between rainfall and flow at the catchment outlet, or of the attenuation effects of the catchment on flow peaks.

The current study aims to explore if the rainfall function can predict the varying pollutant behaviour during an event. One method of achieving this would be to couple *Equation 6.1* with an existing hydrological model and evaluate the results. This however adds another source of uncertainty into the predictions – that of the hydrological model. To avoid this uncertainty, the decision was made to use the observed flow to explore catchment and event characteristics such as lag time and attenuation. These results were then used to examine the rainfall function's performance.

Both lag time and the attenuation are well understood in urban hydrology and have been described many times in the literature (Pilgrim (ed.), 1987). The lag time is the time taken from the rainfall to the time when the runoff is measured. The attenuation is the smoothing of the flow peaks associated with varying rates of flow and storage from different parts of the catchment. The effects of both phenomena on a plot of pollutant flow rate versus time are illustrated in Figure 6.2.

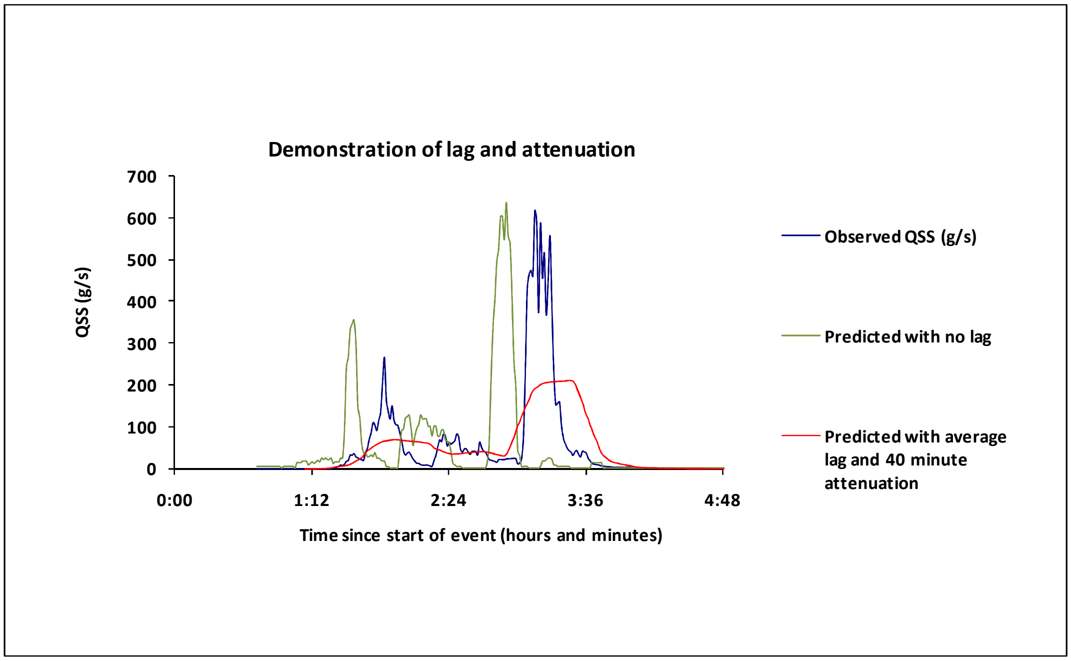


Figure 6-2 Example of the effect of lag time on pollutographs

The first step in describing the effects was to measure the recorded lag time for each event. This was achieved by minimising the sum of least squares (Legates and McCabe, 1999) between the rainfall peaks and runoff peaks for all timesteps from t = 0 (flow recorded at the same time as the rainfall) to t = 65 (runoff recorded 65 minutes after rainfall). This lag time was chosen as it was higher than the time of concentration of any of the catchments when calculated using traditional methods (Pilgrim (ed.) 1987).

Comparison of predicted versus observed behaviour

The comparison of predicted versus observed loads at a timestep was undertaken by comparing the predicted and observed loads at every timestep where a water quality sample was taken. (see Table 6-1 for example – there are 10 samples and therefore 10 points for comparison).

The load over six minutes was calculated by multiplying the concentration by the total flow for the 6 minutes. Given that the model (Equation 6-1) predicts the load in a timestep, the predicted concentration was obtained by dividing the predicted pollutant flow rate by the observed flow.

Table 6-1 Sample data for explanation of within event model assessment

Time	Flow (l/s)	TSS (mg/l)	Observed load (kg in 6 minutes)	Predicted load(kg in 6 minutes)
1:24	84.0	33	1.0	1.0
1:29	131.4	84	4.0	4.2
1:34	413.5	87	13.0	13.6
1:39	314.2	63	7.1	7.5
1:44	499.1	220	39.5	41.5
1:49	1287.2	160	74.1	77.9
1:54	1252.8	120	54.1	56.8
1:59	998.6	68	24.4	25.7
2:04	646.7	49	11.4	12.0
2:09	257.7	38	3.5	3.7

This analysis resulted in a set of observed/predicted points for each event collected (both in terms of pollutant flow rate and concentration). As discussed, two additional factors need to be accounted for in predicting the load; the lag time and the attenuation. The influence of these terms was assessed by trialling different combinations of lag and attenuation times and measuring the correlations (R^2) between predicted and observed loads for each combination. The three lag times trialled were:

- No lag (rainfall and runoff occurring at exactly the same time)
- The catchment mean lag (as determined by averaging all lag times for the catchment)
- The optimum lag for each event (measured by taking the individual lag times for each event as described earlier in this section)

In addition to the lag time, water flowing through a catchment is also attenuated over an event. Hydrological models represent this through a variety of means such as the use of subcatchments to generate flow with different lags (Huber and Dickinson 1988) and the use of routing methods (CRCCH 2003). This effect is particularly important when using the rainfall function as it is a discontinuous series. Hence, the model will only predict a pollutant load in the same timestep in which rainfall occurs

To simulate attenuation in the model, the raw ΣI^b figures (derived from each rainfall tip) were averaged over the attenuation time to produce a continuous time series of pollutant flow rate in g/s (see example in Table 6-2). Increasing attenuation on a pollutograph flattens the line – decreasing the peaks and increasing the troughs (Figure 6-3).

Table 6-2 Example of averaging data to simulate attenuation

Time	ΣI^b	10 minute attenuation ($\Sigma I^b/10$)
0:41		0.12
0:42		0.12
0:43		0.12
0:44	0.076	0.12
0:45		0.12
0:46		0.12
0:47		0.12
0:48		0.12
0:49	1.076	0.12
0:50		0.32
0:51		0.32
0:52		0.32
0:53		0.32
0:54		0.32
0:55	2.0761	0.32
0:56		0.32
0:57		0.32
0:58		0.32

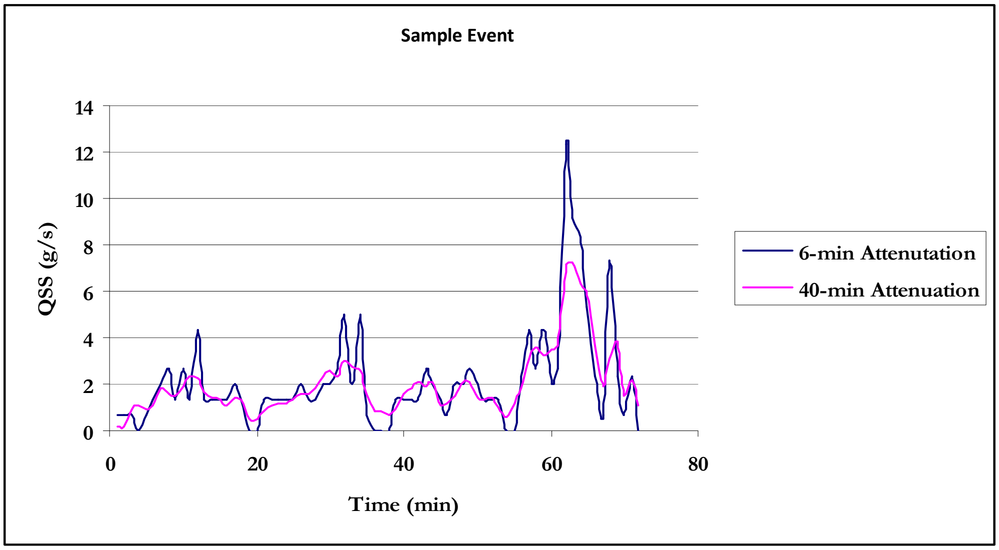


Figure 6-3 Effect of simulated attenuation on the pollutograph (where QSS is the flow rate of pollutants in g/s)

Each event was tested by comparing the three different lag scenarios each with at least three different attenuation values and optimising the results. The coefficient of determination was used to assess the strength of the relationship.

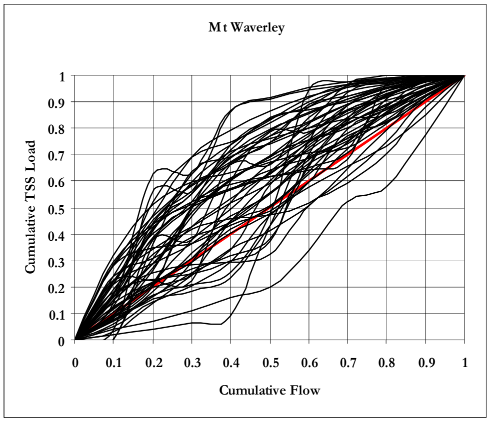
Sensitivity of short term predictions to spatial variability in rainfall

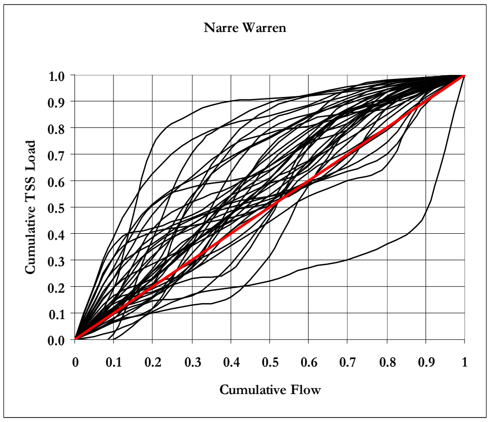
To test the sensitivity of the short term rainfall measurements a test was carried out using two different rainfall gauges close to the monitoring site. TSS was predicted from each of the two gauges surrounding the Richmond site described in Chapter 3. Both rainfall gauges were on the catchment boundary. The rainfall records from each gauge were used to independently predict the pollutant behaviour over a sample event.

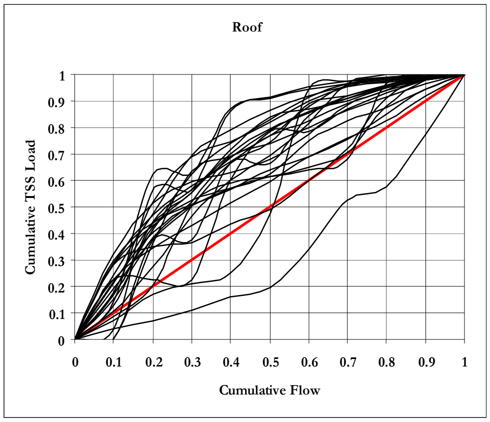
6.4 Results

Description of pollutant behaviour- First flush - cumulative flow versus volume descriptions

The graphs of normalised flow versus pollutant loads (Figure 6-4 for TSS, Appendix D – TP/TN Cumulative Load versus Cumulative Volume Graphs for TP and TN) show that, despite considerable scatter, the majority of events do sit above the 45 degree line, indicating that there is a demonstrable effect of load preceding flow volume.







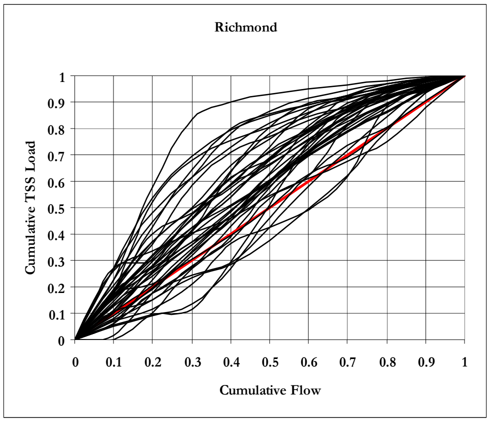


Figure 6 -4 TSS Cumulative Load versus Cumulative Flow Volume Graphs

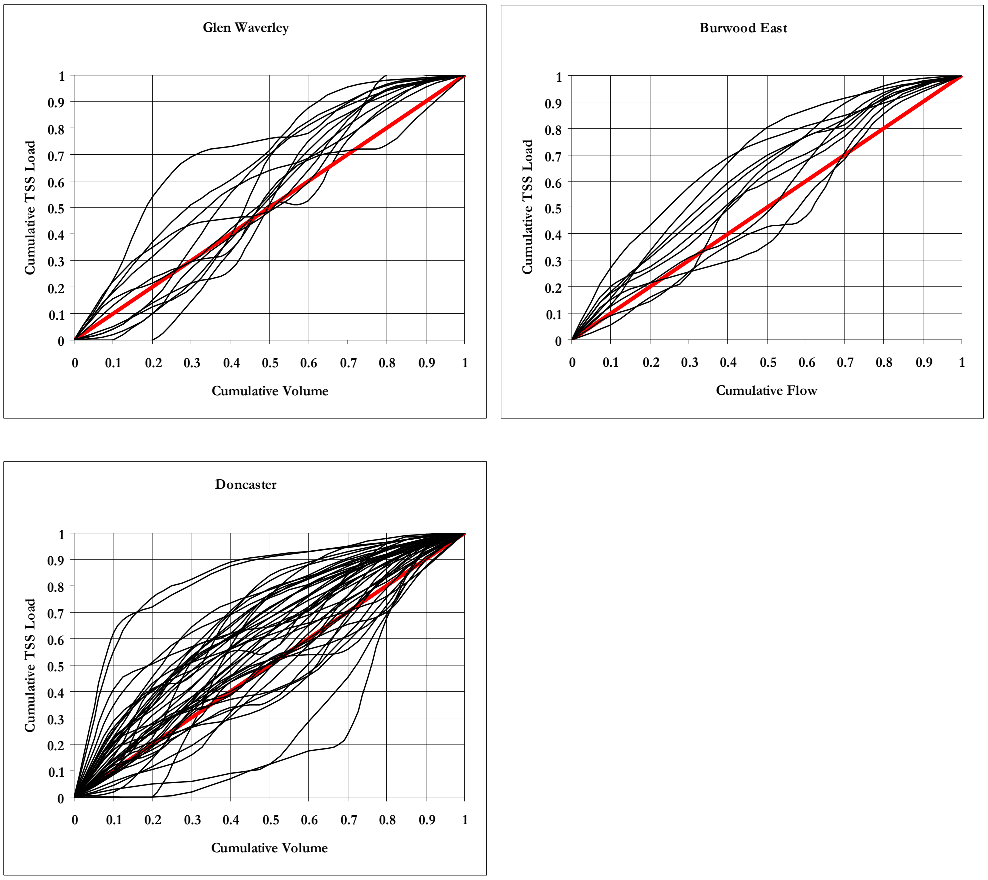


Figure 6-4 TSS Cumulative Load versus Cumulative Volume Graphs

When these results are expressed in a quantified way, however, the effect is not strong. MFF20 values are usually in the 1.2 -1.4 range except for the Monash Roof, where values are between 1.8 and 2.3 (Tables 6-3, 6-4, 6-5). The results for the Roof support the explanation of first flush being a phenomenon which primarily occurs in small catchments (Lee, Bang et al. 2002). Both the graphs (Figure 6.4 for TSS and Appendix D for TP/TN) and the quantitative data show a decreasing strength of the effect from TSS to TP and then to TN

Although the results shown in this type of plot are described as “first flush”, the underlying phenomena require explanation. A distinct observation of load preceding volume is traditionally explained by attributing the first flush to an excess of pollutants being available early in an event, which are washed off as the event goes on. If this were the case we would expect the first flush effect to be greater with larger events as a greater proportion of the load of pollutants would be washed off in large storms, therefore providing more opportunity for depletion. To test this, the analysis was repeated with only large events (above 10mm rainfall; this was undertaken for TSS only as other pollutants had few large events). Results were similar to those presented for all events (Table 6-6). There was no increase in the amount of load carried by the first 20% of the storm.

Given the results of chapter five, where rainfall intensity is shown to be the significant driver behind loads (as opposed to flow volume), an alternative explanation of the results could be that the rainfall intensity peaks before the flow (as rainfall drives flow) and therefore the peak in pollutant loads is in fact a function of the energy supplied by the rain, mobilising pollutants. The lower figures recorded for nitrogen (Table 6-5) could be caused by a relatively large proportion of the nitrogen being a component of the rainfall rather than being generated during surface contact.

Table 6-3 Summary of Cumulative Pollutant Loads (TSS)

TSS	Number of Events	Number of events above 45% line when 20% of the total flow has been recorded	MFF_{20}
Mt Waverley	49	33	1.4
Glen Waverley	19	12	1.2
Narre Warren	38	26	1.5
Monash Roof,	24	22	2.0
Burwood East	10	8	1.4
Richmond	40	29	1.4
Doncaster	42	30	1.5

Table 6-4 Summary of Cumulative Pollutant Loads (TP)

TP	Number of Events	Number of events above 45% line when 20% of the total flow has been recorded	MFF_{20}
Mt Waverley	23	14	1.3
Glen Waverley	12	5	1.2
Monash Roof	13	12	2.3
Richmond	21	19	1.4

Table 6-5 Summary of Cumulative Pollutant Loads (TN)

TN	Number of Events	Number of events above 45% line when 20% of the total flow has been recorded	MFF_{20}
Mt Waverley	23	15	1.2
Glen Waverley	12	6	1.1
Monash Roof	13	12	1.8
Richmond	21	19	1.4

Table 6-6 Summary of Cumulative Pollutant Loads (TSS), events over 10mm rain

TSS (over 10mm rain)	Number of Events	MFF_{20}
Mt Waverley	15	1.4
Glen Waverley	6	1.2
Narre Warren	16	1.5
Monash Roof	3	2.0
Burwood East	6	1.4
Richmond	13	1.4
Doncaster	9	1.5

Description of pollutant behaviour - Depletion of pollutants with time and volume
The observed/predicted load ratio was tested against event duration and cumulative volume. If either of these factors were having a significant impact , we would expect to see distinct slopes on the graphs (see Figure 6-5 for example).

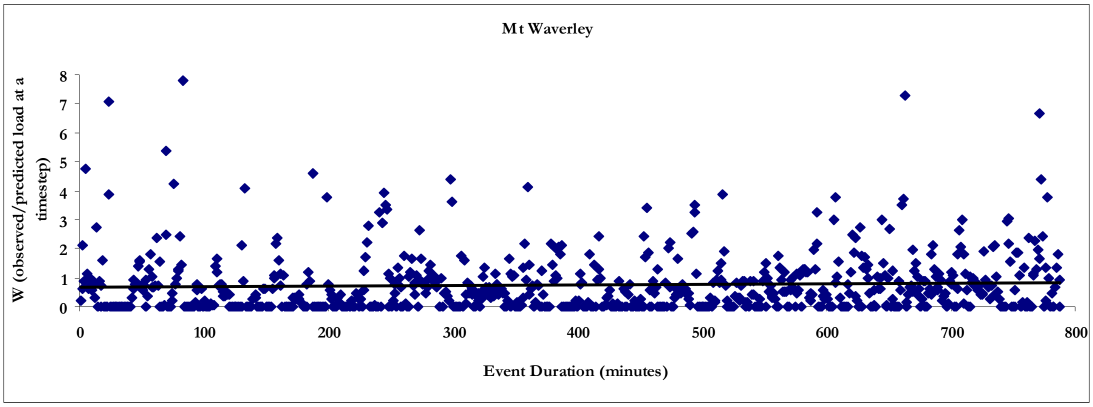


Figure 6-5 Observed/predicted load ratio versus event duration – all measured points at Mt Waverley (TSS)

For duration or volume versus *W* relationships none of the catchments show strong relationships for any pollutant (all R^2 values are less than 0.2 with most less than 0.1).

Richmond is the only catchment with a consistent slope with all three pollutants declining with time and cumulative volume. Overall more combinations show a significant relationship with duration than cumulative volume (Table 6-7), however the scatter and lack of significance is dominant and there is certainly no compelling argument from these observations for a depletion term.

Table 6-7 Relationship between *W* and Duration/Cumulative Volume

	Relationship – W and:				
	No. Observations	**Duration**		**Cumulative Volume**	
TSS		Slope	R^2	Slope	R^2
Mt Waverley	509	-0.0011	0.00	0.0000	0.00
Glen Waverley	251	0.0002	0.00	0.0000	0.00
Narre Warren	147	-0.0004	0.00	-0.0001	0.00
Monash Roof	183	0.0000	0.00	-0.0005	0.00
Burwood East	191	-0.0005	0.00	0.0000	0.00
Richmond	377	-0.0008	0.00	0.0000	0.00
Doncaster	381	0.0006	0.00	0.0000	0.02
TP					
Mt Waverley	299	-0.00031	0.00	0.010505	0.01
Glen Waverley	134	0.00098	0.01	0.000018	0.02
Richmond	185	-0.00386	0.04	-0.000007	0.04
Monash Roof	109	0.00000	0.03	-0.000003	0.08
TN					
Mt Waverley	285	-0.000346	0.01	0.000004	0.01
Glen Waverley	163	0.000013	0.00	0.000013	0.02
Richmond	197	-0.001654	0.05	-0.000003	0.01
Monash Roof	117	-0.000001	0.05	-0.000009	0.11

Comparison of Predicted and Observed Pollutant Behaviour- Calculation of catchment lag times

Lag times were calculated for all events in order to match the predicted pollutagraphs with the observed behaviour. Generally the times calculated are as expected, with the larger catchments having longer times. The scatter in lag times between events for the same catchment is however significant (Table 6-8). A large variation in lag has implications for the predictions of short interval pollutant flux as accurate prediction of loads (at a short timescale) requires accurate representation of lag times. This factor is further explored when predications of pollutant flux are made.

The Doncaster catchment recorded a very low average lag. This is thought to be caused by it being relatively steep and having a relatively short flow distance (when compared to area) from the top of catchment to the outlet.

Table 6-8 Average Lag Times for all catchments

	Mean of Lag (minutes)	σ of Lag times (minutes)
Mt Waverley	18	17
Glen Waverley	7	4
Narre Warren	21	18
Monash Roof	5	4
Burwood East	18	12
Richmond	26	10
Doncaster	11	11

Not surprisingly, given the results for calculated lag times, the main factor influencing the correlations between predicted and observed loads at a short timestep is the value of the lag and attenuation used (Table 6-9). When the calculated (optimum) lag is used for the event, the model is, perhaps unsurprisingly, able to predict the observed timestep loads with high coefficient of determination.

These figures demonstrate that the temporal variation within the event becomes the most important variable in short term prediction, and the difference in correlations between an average lag and an optimum lag demonstrates the sensitivity of predicting pollutants at such short timesteps.

Table 6-9 Narre Warren TSS (loads per timestep)

Observed versus Predicted Points for all recorded events	Mean values of coefficient of determination (R^2)		
Attenuation	**No lag**	**Average Catchment lag (22 minutes)**	**Optimum lag for each event**
6-minutes	0.14	0.27	0.75
10-minutes	0.16	0.28	0.75
20-minutes	0.24	0.37	0.69
40-minutes	0.29	0.38	0.59

All catchments and pollutants follow the same pattern with an optimum attenuation and lag combination serving to improve the model's predictions (column 4 in Table 6-9). Based on the correlations the model is capable of achieving good predictions at a short timestep, particularly for loads (Tables 6-10, 6-11 and 6-12). Using the optimum lag in all cases significantly improves the correlations over those using the average lag, or no lag at all (see Table 6-9 for example).

Table 6-10 TSS correlations for within event prediction for optimum lag for each event

TSS	**Loads**	**Concentration**
Catchment	**Mean values of coefficient of determination, R^2 (optimum time lag for each event)**	
Mt Waverley	0.72	0.46
Glen Waverley	0.47	0.50
Narre Warren	0.64	0.70
Monash Roof,	0.61	0.52
Burwood East	0.65	0.73
Richmond	0.85	0.58
Doncaster	0.49	0.68

Table 6-11 TP correlations for within event prediction for optimum lag for each event

TP	**Loads**	**Concentration**
Catchment	**Coefficient of Determination (R^2) with optimum lag for each event**	**Coefficient of Determination (R^2) with optimum lag for each event**
Mt Waverley	0.64	0.65
Glen Waverley	0.66	0.48
Monash Roof	0.46	0.62
Richmond	0.79	0.74

Table 6-12 TN correlations for within event prediction for optimum lag for each event

TN	**Loads**	**Concentration**
Catchment	**Coefficient of Determination (R^2) with optimum lag for each event**	**Coefficient of Determination (R^2) with optimum lag for each event**
Mt Waverley	0.70	0.56
Glen Waverley	0.65	0.45
Monash Roof	0.77	0.53
Richmond	0.51	0.49

Sensitivity of short term predictions to spatial variability in rainfall

When different rainfall gauges were used (both within the Richmond catchment) to predict the pollutant behaviour, the analysis showed at minimum a 5-minute difference in timing between peaks, an approximate 25% difference in magnitude of the second peak, and an approximately 5% difference in total load (Figure 6-6). This strongly suggests that the spatial and temporal variability of rainfall, even in small urban catchments, is a significant barrier to consistent prediction of short term pollutant behaviour.

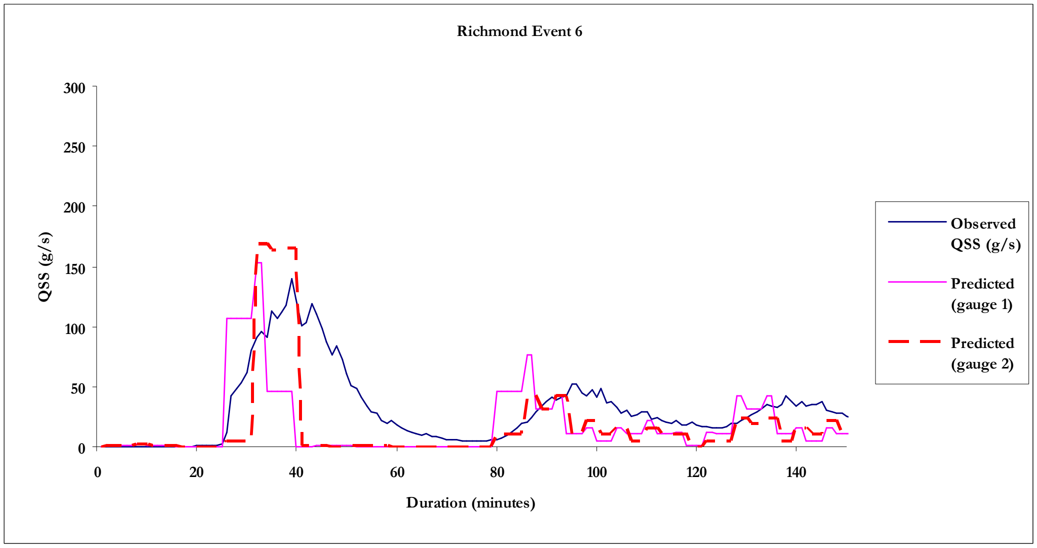


Figure 6-6 Difference in predicted pollutant flow rates (versus observed) when different rainfall gauges were used for prediction (where QSS is the flow rate of pollutants in g/s). Note that even when both rainfall gauges are within the catchment the spatial and temporal variability significantly changes predictions.

6.5 Discussion

Trends in pollutant behaviour

It is apparent from the data that the first flush effect, if present, is not a major influence over behaviour and is not a strong or consistent enough phenomenon in any of the monitored catchments to justify its use in treatment measure design.

The first flush data is consistent with past studies, in that the small roof catchment shows a more distinct effect than the larger catchments. The data is also consistent with an explanation of rainfall intensity driving the pollutant suspension from urban surfaces, and then being carried by the flow to the catchment outlet. This explanation is supported by the presence of a stronger rainfall-intensity effect for TSS, followed by TP and TN, in decreasing order of mean particle size distribution.

The results suggest that rainfall intensity could provide a possible explanation for the ‘first flush’ effect because of the timing of rainfall hitting surfaces and mobilising pollutants (Table 6-9) before the peaks in runoff occur. Further work on small catchments would be required to test this.

Most importantly there was no consistent pattern to the pollutant behaviour over an event. The results shown in Figure 6-1 gave no evidence of depletion, and indeed there were as many trends with positive trends as there were with negative trends. In some cases the positive trend may be explained by the influence of the large events introducing pervious area runoff into the system. In these events pervious area runoff is likely to have been present to such an extent that flow was concentrated over unpaved areas contributing pollutants through disturbance as opposed to rainfall impact on impervious surfaces.

The lack of a strong pattern in pollutant behaviour does however raise two possibilities:

- The analysis undertaken was not sensitive enough to uncover any effects of depletion
- There is little or no depletion occurring in the events monitored

By comparing the observed versus predicted behaviour, more insight will be gained into this, as if high correlations are achieved at short timesteps over an event, it will pose further questions about the need to include a depletion term in predictive models.

Comparison of Predicted and Observed Pollutant Behaviour

The testing indicates that, given an accurate representation of the lag time and attenuation, the method predicts the intra-event (short-duration) pollutant behaviour acceptably. High correlations are achieved for all pollutants if the lag term is optimised to each event. If an average lag term is used the performance for both loads and concentration drops markedly. The distinct differences in correlations between average and optimum lag however flag an important question in the modelling of pollutant flux, especially if the spatial variation in rainfall is causing the impact as suggested by the trial of different rainfall gauges at the Richmond site. It also may explain why previous attempts to achieve this have had limited success.

Indeed the variance in lag times illustrated one of the difficulties in modelling short-term loads (or concentrations). Extremely complex modelling of pipe networks, their slopes and conditions is likely to be necessary to be able to adequately simulate the flow behaviour. Even if this is possible the temporal and spatial variability in rainfall (which is further discussed in Chapter 7) is likely to be a factor in prediction. Given that the results show that peaks can be modelled to a reasonable degree of accuracy, the need for exact timing should be questioned as in most cases there is no management need for prediction at such a short timestep. The need for such data should be carefully considered in modelling design, given the ‘cost’ in terms of data required for calibration.

The results show that the rainfall function can predict the changes in loads and concentrations well over the duration of an event, adding further evidence that the energy contributed by rainfall is the significant driver behind pollutant generation. There is very little evidence to suggest that further terms are needed to describe the pollutant flow, as depletion or first flush are not strong factors in the data presented.

7 IMPACT OF INPUT DATA UNCERTAINTY

7.1 Introduction

Uncertainty is a topic central to both research and practice of stormwater management. The uncertainties associated with hydrological and pollutant modelling are significant (Ahyerre et al. 1998) and this uncertainty has implications for the way stormwater is managed (Bertrand-Krajewski et al. 1993). Modelling without assessing uncertainties of the produced results is no longer considered acceptable.

Based on the discussion in Section 2.8, three general sources of uncertainty have been considered with respect to the main topic of this thesis;

1. Uncertainties due to monitoring (response) data errors
2. Uncertainties due to the model structure
3. Uncertainties due to input data errors

Monitoring or response uncertainty is taken here to represent those errors associated with the collection of data. In the case of the current study this has included errors in the collection of flow and concentration data and the structure of the monitoring program itself. Investigations into sampling and analytical error have returned uncertainties similar to those reported elsewhere in the literature with uncertainties of approximately 30% being associated with EMC data. The primary monitoring program was then designed to capture enough events so that the EMC uncertainty was reduced to a minimum, approaching the experimental uncertainty of the data collection (i.e. enough events were captured to represent the widest range of behavioural response from the catchment). This was only achieved in a limited number of cases; however the initial work done gives an understanding of the uncertainty associated with the monitoring program, which can then be compared to the results of the model.

The uncertainty associated with the model itself (Grayson et. al 1992) concerns factors such as the model structure, the timestep used and the availability of calibration data. These issues have been discussed in Chapter 5, which presents an examination of the model calibration and the sensitivity of model parameters. For example, model parameters (b and W) appear to be correlated with each other in most cases.

The focus of this chapter is the third source of uncertainty - the error associated with model inputs. In order to predict loads the model requires two inputs, being (1) Rainfall intensities, and (2) Catchment area. Clearly, any uncertainties in these data will produce uncertainties in the model results. For example, uncertainty surrounding the spatial and temporal variability of rainfall is a major factor in the reliability of hydrological models (Jakeman and Hornberger 1993; Ogden and Julien 1993). In a study that used Monte-Carlo approaches to examine uncertainties in combined sewer systems Arnberg-Nielson and Harremoes (1996) determined that description of the rainfall provides the greatest uncertainty.

For the current study a third input, lag time, could be added if predictions of short term pollutant flux were required. As the likely outcome of this study however is to improve the prediction of pollutant loads however the error analysis has been undertaken with only those input parameters necessary for load prediction.

The datasets collected for this study have been used to more thoroughly examine the uncertainties in linked hydrological and pollutant generation models (Dotto et al 2009). The work is following the methodology of Kuczera (2006) who recommends that all sources of uncertainties should be considered (and propagated through the model) at the same time, as they can compensate for each other if considered separately. Initial work on subsets of the primary data has shown that input uncertainties (only rainfall data was considered) have impacts on the distribution of the calibration parameters b and W. They reinforce the findings from Chapter 5 which show that the calibration parameters are related.

This chapter examines the sensitivity of the model outputs to changes in rainfall and area inputs. Further studies will continue (Dotto et al 2009) to combine these under a "total error framework", but this approach is beyond the scope of this thesis.

7.2 Method

The analysis in this chapter will concentrate on the experimental uncertainty associated with the input data to the model – namely the rainfall at 6 minute timesteps and the total area of each catchment.

The experimental uncertainty of input data can be considered to be made up of two components, systematic and random errors:

- Systematic error - errors where all the measurements are in a particular direction, too high or too low. In the case of rainfall data this may be due to poor calibration, incorrect clock adjustments, or the behaviour of tipping gauges at varying rainfall intensities that have been reported frequently in the literature (Duchon and Essenberg 2001; LaBarbera et al. 2002; Strangeways 2004). The poor setup of a rainfall gauge at an incorrect height, or too close to nearby trees would also cause systematic errors

- Random error - errors associated with an individual measurement such as the interference of a raingauge by falling debris or a power outage. A major source of random error is the use of a point data to predict a catchment rainfall. The error associated with the calculation of area on the GIS system is taken as being all random

The methodology for compiling errors has therefore been initially split into two sections, systematic and random. These areas have been discussed separately, as the methods for addressing each of the types of errors in practice are different. Systematic errors can be minimised by good experimental design and regular calibration procedures. Random errors are more difficult to address but can be recognised and quantified as an uncertainty inherent in the monitoring procedure.

Systematic Errors

Variations in rainfall and uncertainties around its measurement and representativeness are well studied. For many hydrological and stormwater quality models the rainfall input has been demonstrated to be the largest source of error in prediction (Bedient, Lambert et al. 1980; Ahyerre, Chebbo et al. 1998).

Systematic errors from tipping bucket rain gauges can be caused by poor calibration or from physical factors such as outsplash and wind effects. Studies from the United Kingdom demonstrate that an elevated raingauge (such as those used in the current study) typically under-predict total rainfall by approximately 6% (Strangeways 2004). This effect increases for large rainfall intensities to 9 – 15% (Marsalek 1981; LaBarbera et al. 2002) although some authors have reported higher uncertainties for long duration, high intensity events (Molini et al. 2005).

Many of these uncertainties can be reduced through the use of dynamic calibration (calibration in which the input varies over a specific length of time and the output is recorded versus time), however this is rarely used in practice, particularly with the 6-minute data that is likely to be used in water quality models.

Uniform Systematic Errors

The model tested in chapter 5 (**Error! Reference source not found.**) was used to test the impact of systematic errors on load predictions. It was assumed that there were no systematic errors in the calculation of area, only in the rainfall measurement.

Equation 5-1 was run for seven scenarios using rainfall that was adjusted at each tip (–20, -10, -5, 0, +5, +10 and +20% errors). For instance in the -20% trial the rainfall total for each 6 minute timestep was adjusted down by 20% before being subject to the power term. The area term (A), power (b) and calibration co-efficient (W) used were the optimum values obtained for a 6-minute calibration obtained in Chapter Five. It is expected that this will give a consistent percentage error for all events which will be dependent on the power term "b". A higher "b" would increase the impact of the adjustment on each rainfall tip. The results were tabulated for each pollutant (all sites combined).

Systematic error dependent on rainfall intensity

Secondly, to trial the impact of potential under representation of rain gauges at high rainfall intensities a second run was trialled based on a staged approach (Marsalek 1981; LaBarbera, Lanza et al. 2002; Molini, Lanza et al. 2005). The approach proposed is:

$$I_i = \begin{pmatrix} I \longrightarrow 0 < I < 20mm/hr \\ 1.1 \times I \longrightarrow 20mm/hr < I < 80mm/hr \\ 1.15 \times I \longrightarrow 80mm/hr < I \end{pmatrix} \qquad \textbf{Equation 7-1}$$

where *I* is the recorded rainfall and I_i is the adjusted rainfall.

Each event was run for the same seven scenarios described above, however the rainfall intensity input was adjusted using the "staged" approach of $I_i = \begin{pmatrix} I \longrightarrow 0 < I < 20mm/hr \\ 1.1 \times I \longrightarrow 20mm/hr < I < 80mm/hr \\ 1.15 \times I \longrightarrow 80mm/hr < I \end{pmatrix}$ Equation 7-1 rather than a straight percentage adjustment.

This approach aims to examine the possible impact of greater uncertainties at very high rainfall intensities. The factors used in the staged approach are the maximum recorded in the literature; therefore the results are likely to be conservative. To examine these results a sensitivity analysis was undertaken by constructing plots of the relative uncertainty of each scenario versus the systematic error (-20%, -10% and so on).

To assess how likely this scenario is a simple comparison was undertaken between the runoff co-efficient for all events that had rainfall intensities above 20mm/hour and those with maximum rainfall intensities below this. If the gauges were underestimating the rainfall at high intensities it is hypothesised that the average runoff co-efficient would be higher for events where rainfall intensity is high.

Random errors

Both rainfall and catchment area could contain random errors. The Law of Propagation of Uncertainties (Talylor 1994) was used to calculate the relative error in the load predictions by assuming an error for both the rainfall and area terms in Equation 5 -1.

Rainfall

In even small urban catchments the spatial variability of rainfall has been shown to significantly affect the total volumes and the intensity reported within an event (Berndtsson and Niemczynowicz 1988; Nelen et al. 1992; Chaubey et al. 1999; Moreira et al. 2006). Absolute errors of 30% have been recorded (Molini and Lanza, 2005) for raised tipping bucket gauges. For the purposes of this study these differences are considered to be random, for each particular timestep it is unknown how representative the gauge is for the catchment. Likewise over a total event it is unknown if the gauge is under or over predicting the catchment rainfall and it is unlikely there is a consistent pattern between events although some authors have reported that a single gauge will over predict in large and under predict in small events (Troutman 1983). An uncertainty of +/- 30% (uniform distribution) has been adopted for each tip during an event.

Area

As described in section3-8 the area was calculated using local government drainage plans, aerial photography, and contours on a computer based GIS system. Field trips to the catchments were also used to verify the drainage layouts.

To test the uncertainty, four experienced GIS users (from Melbourne Water's Waterways Group) were asked to measure the catchment area of two catchments using the same information (aerial photographs, local underground drainage plans, 1m contours and road layouts). The four estimates were then compared to determine the percentage error in the area estimation (Table 7-1). An uncertainty of +/- 10% (uniform distribution) has been adopted for each area calculation.

Table 7-1 Area calculations for Doncaster and Mount Waverley catchments

	Area (Doncaster, ha)	**% difference**	**Area (Mount Waverley, ha)**	**% difference**
Original Calculation (including ground truthing)	105.7		28.0	
GIS User 1	110.8	5	26.2	6
GIS User 2	101.2	4	29.1	3
GIS User 3	97.6	8	26.6	5
GIS User 4	104.3	1	30.9	10

Propagation of Uncertainties

The combined data from these trials was used to determine an uncertainty associated with the model by using the Law of Propagation of Uncertainties shown in Equation 7.2 (Talylor 1994).

$$u(y)^2 = \sum_{i-1}^{N} u(x_i)^2 \left(\frac{\delta f}{\delta x_i} \right)^2 + 2 \sum_{i=i}^{N-1} \sum_{j=i+1}^{N} R(x_i, x_j) u(x_i) u(x_j) \frac{\delta f}{\delta x_i} \frac{\delta f}{\delta x_j}$$ **Equation 7.2**

where $R(x_i,x_j)$ is the coefficient of correlation between the quantities x_i and x_j. Equation 7.2 is based on a first-order Taylor series approximation of Y = f(X1,X2,…,XN).

The equation is based on a Y that is determined as a function of N factors that may be measured such that Y = f(X1,X2,…,XN). As Xi and Y are not known values they are estimated as xi and y (xi is usually the mean of a number of trials). The uncertainty of xi (u(xi)) is given by the standard deviation of the mean. The true value of the estimate xi has approximately 95% probability of being in the range xi±2u(xi). Equation 7.4 calculates the uncertainty of y (u(y)).

The relative uncertainty of a quantity xi (Δxi/xi), and similarly for y, can be expressed as:

$$\frac{\Delta x_i}{x_i} \approx \frac{2u(x_i)}{x_i}$$ **Equation 7.3**

The Law of Propagation of Uncertainties was used to calculate the relative uncertainty for each event (and pollutant) for each catchment. Equation 5-1 was used, with the previously calibrated values of b and W at a 6 minute timestep. When the rainfall function equation is applied:

$$Load \; = \Delta t W A \sum I^b = \Delta t W A (I_1^b + I_2^b + I_3^b + I_4^b ....)$$

We assume that $u(\Delta t) = 0$, therefore

$$u(L^2)=u(A)^2\left(\frac{\delta L}{\delta A}\right)^2+u(I_1)^2\left(\frac{\delta L}{\delta I_1}\right)^2+u(I_2)^2\left(\frac{\delta L}{\delta I_2}\right)^2+u(I_3)^2\left(\frac{\delta L}{\delta I_3}\right)^2+....$$

$$+2R(I,I_2)\mu(I_1)\mu(I_2)\frac{\delta L}{\delta I_1}\frac{\delta L}{\delta I_2}+2R(I_2,I_3)\mu(I_2)\mu(I_3)\frac{\delta L}{\delta I_2}\frac{\delta L}{\delta I_3}+....$$

Assuming no correlation between the area (A) and any rainfall term (I) the equation can be simplified as:

$$u(L^2)=u(A)^2\left(\frac{\delta L}{\delta A}\right)^2+u(I_1)^2\left(\frac{\delta L}{\delta_1}\right)^2+u(I_2)^2\left(\frac{\delta L}{\delta_2}\right)^2+\sum_{i=0}^{N}\sum_{j=i+1}^{\min(N,m)}R(I_i,I_j)u(I_i)u(I_j)\frac{\delta L}{\delta_i}\frac{\delta L}{\delta j}$$

Equation 7.4

where m is a variogram function which calculations the significance of the correlation between rainfall timesteps (Equation 7-4).

When the partial derivatives are calculated:

$$\frac{\delta L}{\delta A}=\Delta tW\sum I^b$$

$$\frac{\delta L}{\delta I_i}=b\Delta tWAI_i^{b-1}$$

The terms are inserted and the equation can be expressed as:

$$u(L^2)=u(A)^2\left(\Delta tWAI^b\right)^2+\sum_{i=0}^{N}(I_i)^2\left(b\Delta tWAI^b\right)^2+$$

$$\sum_{i=0}^{N}\sum_{j=i+1}^{\min(N,m)}R(I_i,I_j)u(I_i)u(I_j)\left(\Delta tbWAI_i^{\,b-1}\right)\left(\Delta tbWAI_j^{\,b-1}\right)$$

Equation 7-5

Where a significance correlation in rainfall tips is assumed up until m timesteps, then R = 0.

The variogram function is calculated as:

$$e(t)=\frac{1}{2(M-t)}\sum_{k=1}^{M-t}\left(I_{t+k}-I_t\right)^2$$ **Equation 7-6**

Where the M is the number of values in the rainfall data set and t is the number of timesteps. The equation determines the number of timesteps when e(t) is equal to the variance of the M values of I_t. The variance of the M values is calculated using the var() function in Microsoft Excel 2007.

Two runs were undertaken:

- Using both the rainfall tip uncertainties and area through Equation 7-2 to derive the random uncertainty in the load calculation
- Using only the rainfall tip uncertainty, to compare the impacts of each of the two sources of random uncertainty in the model

7.3 Results

Systematic Errors

As expected, the results of the systematic errors trial in which a constant error was applied to each rainfall tip were purely a function of the scenario tested. Predicted pollutant load varied according to the magnitude of the systematic applied (Table 7-2). All events were consistent for a particular pollutant. The power term in Equation 7 -1 (1.5 for TSS, 1.2 for TP and 1 for TN) meant that the impact was greatest for TSS and smallest for TN.

Table 7-2 Effects of systematic errors on predicted pollutant load (Systematic Errors of -20, -10, -5, 0, 5, 10 and 20%)

Assumed Percentage error (%)	**-20%**	**-10%**	**-5%**	**0**	**5%**	**10%**	**20%**
All Sites							
TSS	-28	-15	-7	0	8	15	31
TP	-25	-13	-6	0	7	13	27
TN	-20	-10	-5	0	5	10	20

When the staged approach (using different values for I based on ranges of intensity, Equation 7-2) was used the function had the effect of greater under representation of the rainfall and therefore the pollutant load. This is demonstrated in Figures **Figure 7-1** toFigure 7-3 (Glen Waverley) with the majority of events (14 out of 19) being represented by the heavy black line. The remaining 5 events all had rainfall intensities that were above 20mm/hr and therefore gave greater under-representation of the rainfall if no other systematic errors were present. All graphs, from all pollutants, show the same pattern with high rainfall intensities being under-represented.

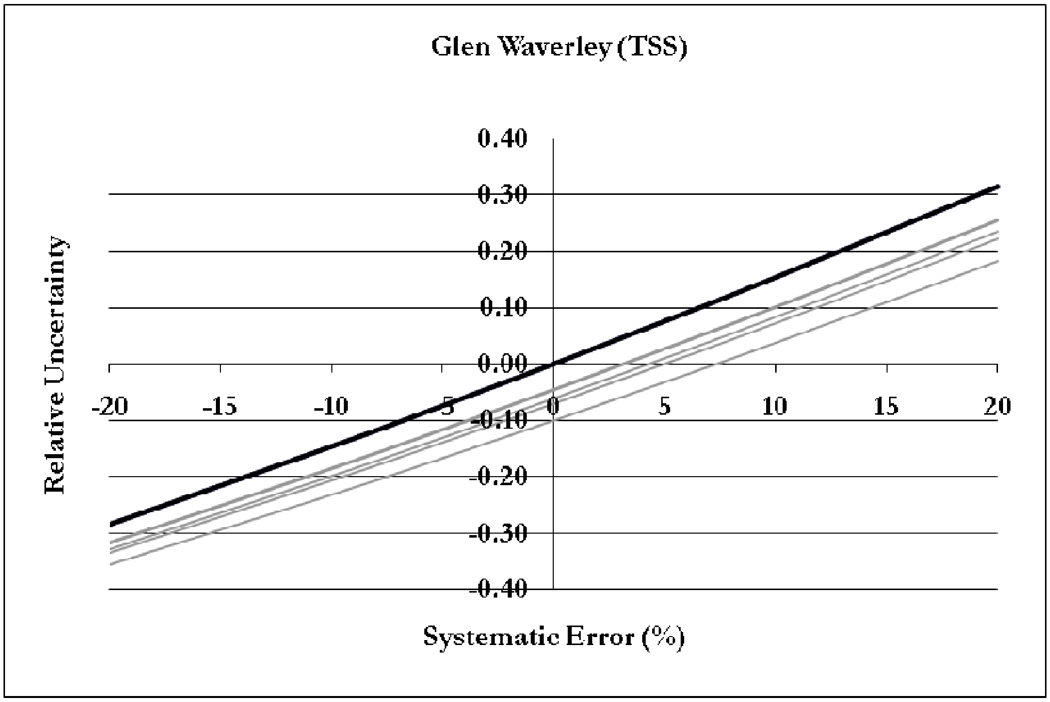


Figure 7-1 Systematic Errors for TSS at Glen Waverley using "staged" approach (the heavy black line represents the majority of events, 14 out of 19)

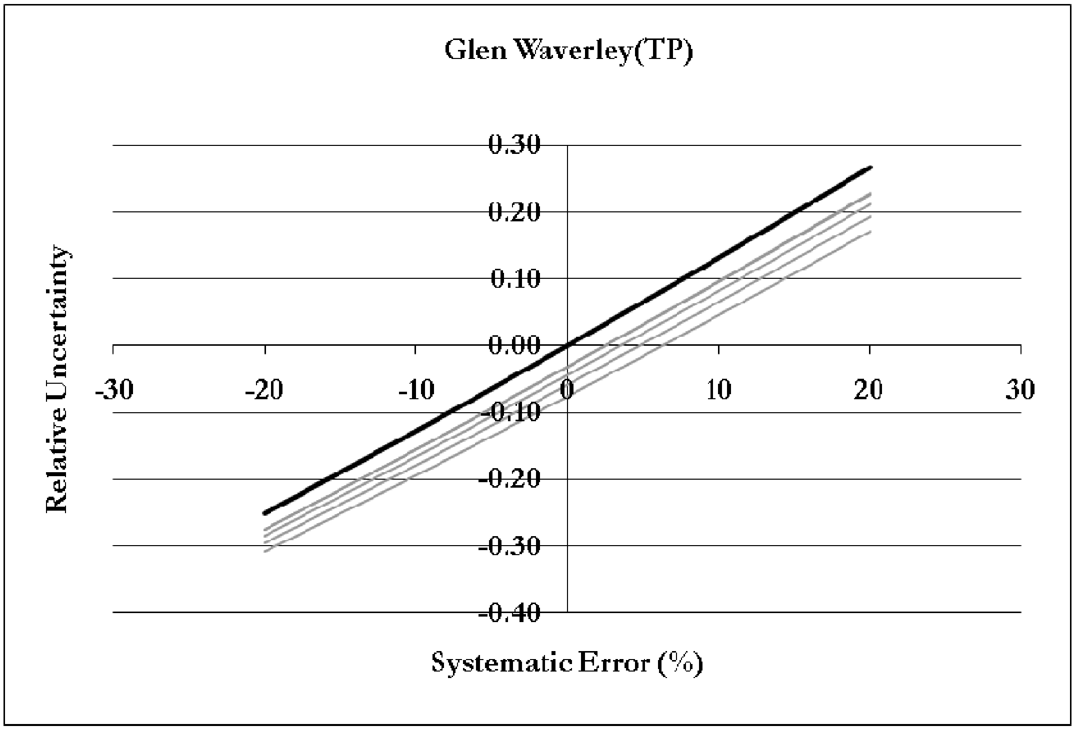


Figure 7-2 Systematic Errors for TP at Glen Waverley using “staged” approach

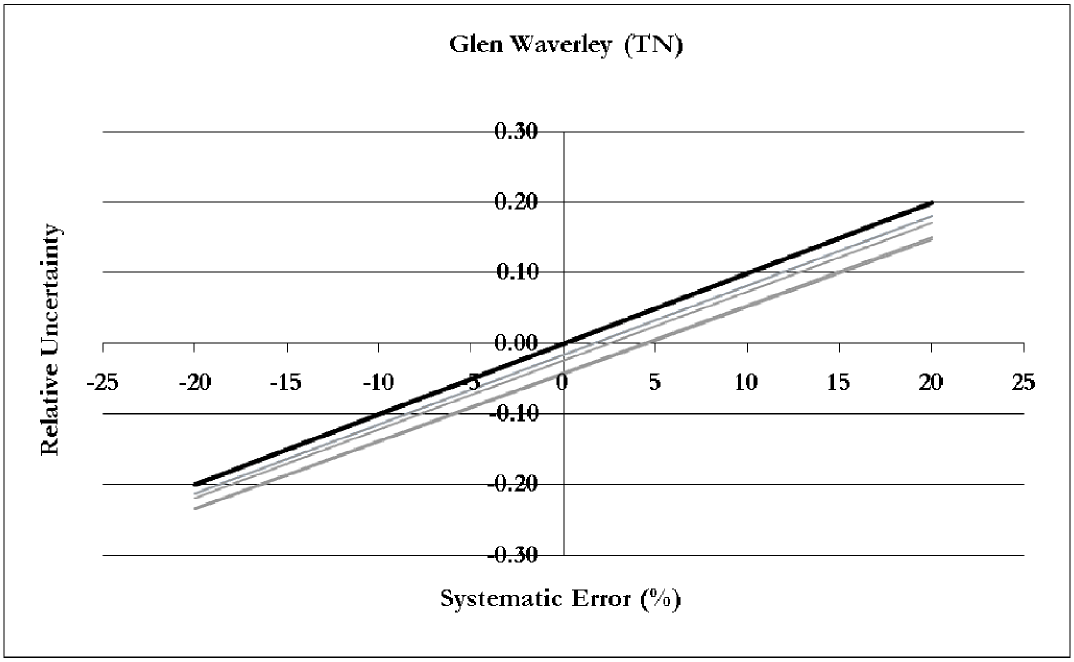


Figure 7-3 Systematic Errors for TN at Glen Waverley using “staged” approach

When a simple comparison of runoff coefficients depending on rainfall intensity events was undertaken (events with maximum intensity less than 20mm/hr, represented by the heavy black line on **Figure 7-1**, Figure 7-2 andFigure 7-3), the average coefficient was lower for high intensity events (Table 7-3) in all seven cases.

If rainfall was being uniformly and accurately measured across the catchment it may have been expected that the low intensity events would have a lower runoff co-efficient. If the rainfall gauges were indeed under representing rainfall however this is the result that would be expected, as lower runoff coefficient could be caused by less rainfall being recorded in the gauge. This hypothesis is based on small numbers of events, however, and does not prove a causal link therefore it should be treated with caution. More extensive analysis would be required, with a greater number of events, to further examine this effect.

Table 7-3 Comparison of runoff coefficients with high and low rainfall intensities

Site	Average runoff co-efficient for events with maximum rainfall intensities > 20mm/hr (number of events)	Average runoff co-efficient for events with maximum rainfall intensities < 20mm/hr (number of events)
Mt. Waverley	0.53 (3)	0.62 (46)
Burwood East	0.21(3)	0.30 (26)
Glen Waverley	0.22(5)	0.30 (14)
Narre Warren Sth	0.19(4)	0.24(34)
Monash roof, Clayton	0.72(2)	0.80(22)
Richmond	0.25(3)	0.27(37)
Doncaster	0.28(2)	0.29(41)

Random Errors

The random errors for all events at each site are presented as the relative uncertainty range (95th percentile range) and the lines showing the maximum and minimum values (

Figure 7-4,

Figure 7-5

Figure 7-6). Median values are shown by the thick red lines.

Median values for the sites are typically 20 –35% for TSS, 20 –30% for TP and 15 – 20% for TN. Both the 95th percentile range and the difference between median values are the largest for TSS. This result is to be expected as the propagation of uncertainty through the model includes the power term *b* from Equation 5.1. A higher *b* value means that the intensity term has a greater influence on the error as it is propagated through the model (Equation 7.2).

The random uncertainty calculated is largely made up of the error contribution from the rainfall term. When the analysis is undertaken without using the error associated with the area calculation there is a relatively small difference in results (Table 7-4). The uncertainty contributed from the area term is less than 10% of the total error in most cases (TN at Mt Waverley, Burwood East and Richmond are the exception, although the rainfall input is still the largest source of error (Table 7-4).

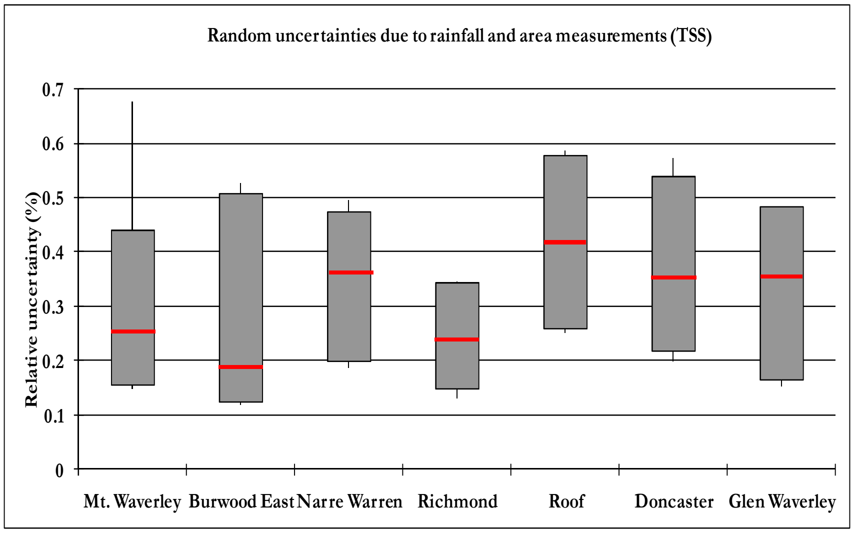


Figure 7-4 Relative uncertainty, TSS all sites (legend shown on Figure 7-5)

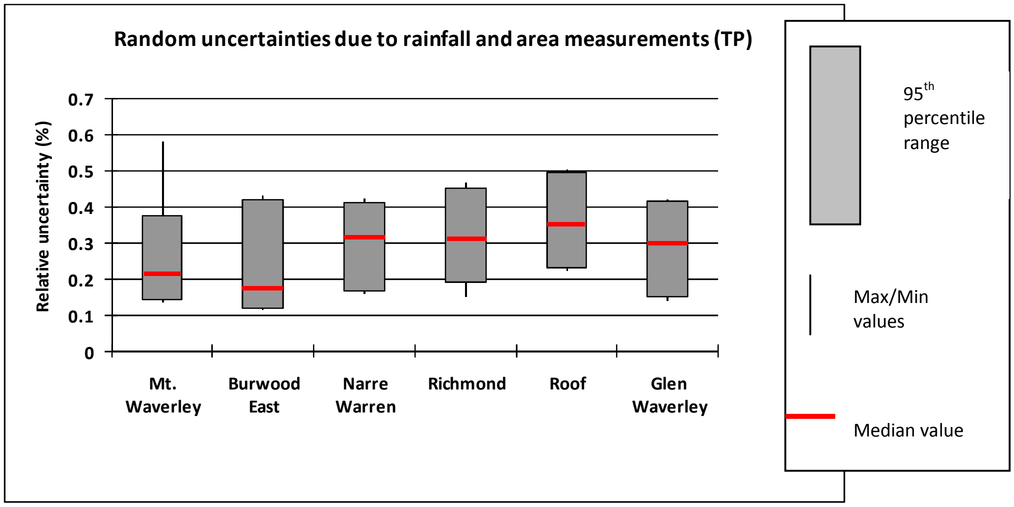


Figure 7-5 Relative uncertainty, TP all sites

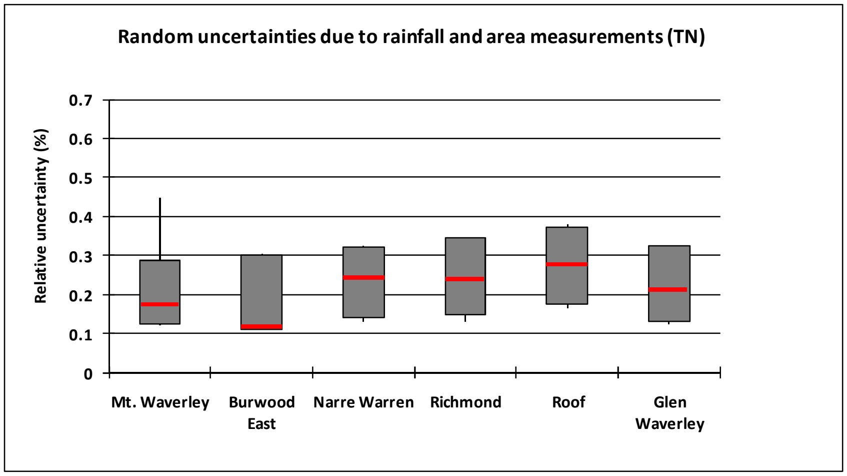


Figure 7-6 Relative uncertainty, TN all sites

Table 7-4 Difference between random error with and without area term (average values for sites)

	Mean Relative uncertainty (%)		
	Pollutant	**Rainfall and Area uncertainty**	**Rainfall uncertainty only**
Mt Waverley	TSS TP TN	27 23 19	25 21 15
Burwood East	TSS TP TN	23 21 17	20 18 13
Glen Waverley	TSS TP TN	35 30 23	33 28 21
Narre Warren Sth	TSS TP TN	36 31 24	34 29 21
Monash roof, Clayton	TSS TP TN	41 36 27	40 34 25
Richmond	TSS TP TN	36 31 24	34 29 21
Doncaster	TSS	37	35

7.4 Discussion and conclusions

Systematic Errors

The results from the analysis of systematic errors was not surprising, with a greater assumed error giving a greater resulting error in model outputs. Because of the power term b, the resulting uncertainty is greater for TSS than for TP and TN. Propagation of systematic errors through the model showed that the impact on pollutant load was about 1:1.5 for TSS– a 10% error in rainfall produced a 15% impact on load. This reduced to 1:1.3 for TP and 1:1 for TN – the impact being directly related to the power term *b*. It must be noted that these types of systematic errors should be reduced through a well set up and maintained monitoring program. All rainfall totals have been compared to nearby gauges (operated by a different organisation), therefore minimising the more extreme variations in calibration in terms of total rainfall.

The staged approach, where the uncertainty was varied according to the rainfall intensity, showed that, if the estimates of systematic error proposed at high rainfall intensities were accurate, pollutant loads could be under-estimated. The comparison of runoff coefficients for high intensity events suggests that these types of errors are possible with the rainfall gauges used for the study. Based on the published estimates it suggests that the model may be underestimating by 5 – 10 % (for TSS) when rainfall intensities increase above 20mm/hr. This effect is also magnified by the power (b) in equation 7-1 with TSS likely to show the largest effect.

Random Errors

The analysis shows that, for load prediction, the rainfall term is the overwhelming source of error in the model. Errors in the area calculation are less important for the load prediction and make up less than 10% of the relative error in all cases. The temporal and spatial variations that occur in rainfall are the main limitation (or constraint) in better predictive ability of event loads.

The random errors show the same pattern as the systematic error analysis, with the TSS load calculation giving higher relative uncertainties than TP or TN. The nature of Equation 5.1, with the *b* term being higher for TSS than TP and TN, results in a higher uncertainty for TSS. Most of the sites generated similar median values in terms of the impact on relative error, with slightly higher figures for the roof site than the other more mixed land use sites. Relative uncertainties of at least 30% are present for all model predictions based on input data. The nature of these errors is extremely dependent on the rainfall monitoring.

The magnitude of the relative uncertainty is comparable to the 95th percentile ranges for EMC's calculated in Chapter 3. The comparison between the variation in monitored EMC's and the relative uncertainty in model outputs gives some confidence that the model results can be legitimately compared with the monitoring data.

8 CONCLUSIONS

The main objectives of the study, as stated in Chapter 2 were to:

1. Review existing relevant data sets on stormwater quality and conduct a large scale monitoring of stormwater in Melbourne
2. Develop a simple predictive model for event loads
3. Test the simple model for prediction of loads and concentrations at short timesteps
4. Undertake sensitivity analyses of model parameters

In this Chapter descriptions of achievements and challenges towards these four objectives are given, together with suggestions for future research and learnings from the current study.

8.1 Review of existing data and monitoring program

A detailed review and exploration of nine existing data sets (both Australian and International) containing over 25 individual catchments was undertaken. Although the data sets had been gathered for various different reasons several factors stood out:

- Data management (despite the best intentions in some cases) was problematic in almost all cases making understanding, quality assessment and use of the data problematical.
- In most cases the data sets had been gathered for management purposes (for instance to establish SMCs) which limited their use for more detailed studies. This was particularly the case when limited pollutant samples had been taken during events.
- In many cases rainfall data (of sufficient quality) was not available within the vicinity of the catchment being tested.
- Uncertainty analysis of the monitoring program had rarely been undertaken.

In the review it was found that, for the purposes of testing the main hypothesis, these data sets did not have the appropriate mix of rainfall and pollutant monitoring – particularly with respect to the temporal aspect of the rainfall measurement. They were also spread over Australia and internationally, making comparison between sites with similar climatic conditions difficult.

As a consequence, the monitoring program undertaken in Melbourne aimed to use consistent methodology and quality assurance in order to overcome these issues. Seven sites were monitored during the program, with approximately 250 events recorded. Only one site obtained over 50 events for TSS (the required number calculated to reduce uncertainty related to variability in EMCs below 30%), but a further three sites recorded over 40 events. Three sites recorded over 30 events for TP and TN, with a further three sites recording between 15 and 20 events. This data set provides an excellent baseline for the understanding of pollutants within the region. In addition, data was collected for metals and other ionic components that had previously not been recorded in any detail within Melbourne. Importantly the study also recorded dry weather flow at several of the site which had higher baseflow concentrations of TN than previously recorded in urban systems. Some of the systems had obvious inputs from septic systems, however the high nitrogen concentrations were uniform across the catchments and should be considered typical. The EMC data recorded was also significantly different to that previous recorded in Australia (TSS was approximately half the concentration recorded in other studies). This follows a similar pattern to recent studies internationally.

The quality of the rainfall and flow data associated with the program means that it will be possible to use it well beyond the current study. Indeed the sites established have been used for further studies on bacterial contamination of urban stormwater (McCarthy 2008) and work into the uncertainty associated with common models (Dotto et al 2009).

As discussed in the description of the monitoring program, the number of events collected fell short of the optimum number of events calculated in Section 3.5 in several cases. If data sets were available with a very large number of events it would be worthwhile to test the hypothesis of how the variability in EMCs behaves with a greater number of events.

Another major area of improvement should be the collection of metals data for the regional context. Only three sites collected metals data, one of those being the roof site so there is not enough variability to make significant conclusions about metals concentrations. Given the increasing focus on the area of stormwater metal concentrations, the driving factors in their generation and their impact on local streams, further data collection would be beneficial.

8.2 Develop a simple predictive model for event loads

Based on the previous work of Duncan (1995), Vaze (2003) and others a methodology was developed and tested based on the short interval rainfall intensity to predict pollution loads. The main hypotheses tested were:

- That rainfall intensity is the primary driver for urban stormwater pollutant loads
- That high quality data is necessary to reduce uncertainties to a scale where the relationship becomes apparent.

The model formulation developed was:

$$Load = W \times A \times T \times \sum_{i=1}^{n} I^{b} \qquad \text{Equation 8-1}$$

where *Load* is the load of pollutant produced by the whole catchment over the period of interest (units of mass), *W* is a calibration coefficient dependent on pollutant and location, *A* is assumed to be the total catchment area (units of length^2), *T* is the duration of the timestep (units of time), *I* is the rainfall intensity in the timestep (units of depth), *b* is a constant dependant on the pollutant, and *n* is the number of time steps over the period of interest.

When tested against other explanatory variables such as total rainfall, total runoff, and a function of the runoff at short timesteps, Equation 8.1 provides clearly better correlations with observed loads. When calibrated against individual catchments, the method returns coefficient of efficiency (E) values of greater than 0.6 for TSS for all catchments except one (0.58 for Glen Waverley). For TP and TN the E values are still above 0.6 in most cases, however also contain some low values (i.e. -0.04 for TP at Narre Warren). When long term loads are considered, the model can predict any of the pollutants within 15% of the measured load, and in most cases within 10%. An exercise was undertaken with all of the large urban catchments (excluding the Roof) to establish if the model still performed well when calibrated with independent data. In all cases (except Narre Warren) the long term TSS load was within 30% of the observed value, and the correlations for event loads also were above E =0.6 in all but one case for TSS. TP and TN still returned promising but less conclusive results with the model failing to predict event loads well in some cases (returned negative E value in four out of eight data sets). It must be noted that the amount of data to calibrate TP and TN models was much less than that for TSS.

These results support the hypothesis that rainfall intensity is a significant driver of pollutant loads generated from urban catchments. While the total amount of rainfall (or runoff) can explain approximately 60% of the variation in loads, the use of a rainfall intensity term at short timesteps, summed over an event can explain upwards of 90%. Given this, the b term in Equation 8-1 becomes important. A b value of one means that the total rainfall is the best explanatory factor of loads. As b increases the role of intensity (and therefore kinetic energy through rainfall impact) is increasingly important. As we would expect b increases with the proportion of particulate matter of the pollutant, with values of 1.5 being calculated for TSS, 1.3 for TP and 1.0 for TN (for a six minute rainfall timestep).

The relationship between timestep and estimated loads is an important finding in the study. Many previous investigations into this area have used averaged rainfall data, or a maximum intensity for an event. The relationship between rainfall timestep, tip size of the gauge and catchment area appears to be important for the precision of the results. If the timestep (as in the case in Belgrade and Lund) is very small the rainfall intensity tends to give binary results – there is little opportunity for several tips to happen within a timestep. In these cases combining timesteps (going from 15 seconds to 1 or 6 minutes) improves the results. The converse to this is if the tip size is large (for instance 1mm in the case of Sandy Creek) it takes very intense rain to produce definition in rainfall intensity. Fortunately, in the case of Sandy Creek there is very intense rain as it is in the sub-tropical climate of south-east Queensland, so the results are instructive. The results suggest that tip size, timestep, catchment area and climatic conditions all need to be considered. The difficulty in finding the most appropriate mix of these factors, particularly when applied to existing data sets such as the secondary sets examined here, is a major reason for the lack of consistent answers to the question of what is the primary driver of urban pollutant loads. Only when a detailed set of data, with repetition between similar catchments, was available, did these factors become apparent.

The results for Melbourne suggest that a 0.2mm tip size and a 6 minute timestep is an appropriate balance for the catchment areas trialled. In some cases the one-minute timestep improved results but not markedly. The results for the Roof catchment were inconclusive with some suggestion that the six-minute timestep was too large, however there was not enough data to fully explore this. Further work should be undertaken on small catchments to examine this relationship.

8.3 Prediction of loads and concentrations at short timesteps

The results achieved when applying Equation 8-1 to event loads give strong evidence that rainfall intensity is the primary driver of pollutant mobilisation in urban catchments, at least in the Melbourne context. If this is indeed the case the model should also be able to predict pollutant behaviour at a short timestep. This is important for two reasons; firstly because a link between intensity and behaviour within events would lead to greater understanding of the processes involved, thus allowing us to further examine stormwater pollution at a fine scale. Secondly, the ability to predict behaviour within an event, to produce pollutographs, may help to improve the management of stormwater pollution. The second reason has been questioned by many authors, and indeed it does not have an immediately applicable use in practice. Satisfactory representation of behaviour however at a fine scale may have uses such as the continuous modelling of pollutant concentrations that are only now starting to be explored in practice.

Because Equation 8-1 is only predicting pollutant loads over a timestep, extra parameters are needed to model concentration flux. Two methods were considered to achieve this; coupling the model to an existing rainfall/runoff model which predicts flow, or using the collected flow data and dividing the load by the flow to obtain concentration. The second method was chosen as it reduced the extra uncertainty added into the process (a rainfall/runoff model) and allowed the best possible examination of the intensity term itself.

The rainfall intensity method also did not have any routing function so that effects such as attenuation and lag time were not immediately predicted. To address this, flow data was used to calculate an optimum lag time and attenuation for each event, which was then used in prediction. This was one of the main findings in this section –when using the rainfall and flow data to calculate lag times there was significant variation between events at the same catchment. This finding greatly affected model performance. If an average lag for the catchment (ie. across all events) was used (such as would be built into many rainfall/runoff models) the model achieved only modest results in predicting loads or concentrations at a six minute timestep (R^2 of 0.2 - 0.5). However, when the optimised lag for individual events was used for the prediction the correlations were high (R^2 of 0.6 - 0.8). Prediction of concentrations using this method was slightly less satisfactory than loads (R^2 of 0.4 - 0.7). There was little difference between the predictions of different pollutants. It is suggested that the spatial and temporal variability in rainfall, even in these small urban catchments, becomes a limiting factor when trying to predict behaviour at a short timestep (in this case 6-minutes).

These results contribute to the understanding of rainfall intensity as a driver of pollutant behaviour, as the model can explain the “peaks and troughs’ of pollutant flux. They do however call into question attempts to model short term pollutant flux. Any progress in this area is likely to be confronted by significant complexity about dynamically predicting rainfall behaviour and therefore the added complexity would need to be balanced with outcomes for the model.

8.4 Sensitivity analyses of model parameters

Understanding of the model structure and the uncertainty associated with the input parameters allows us to gain an understanding of the model’s performance and how it is likely to behave in certain situations. In the case of Equation 8.1, the model is calibrated in an unusual way; first the b term is optimised using regression analysis which then leads to a W value. Given that this method has been used, it is highly likely that these parameters work together as a set. Although it was not apparent in the current study this does indicate that various combinations of parameters may provide a solution set. Preliminary investigation into these topics was undertaken in the current study and more analysis is currently underway using new techniques to examine all uncertainties simultaneously (Dotto et al 2009).

When examined, in most cases, the calibration parameters b and W were correlated with each other. When a wide range of rainfall events was used in calibration, the solution set was far more limited (appeared as a ‘peak” , see Mt Waverley TSS or Richmond TSS in Appendix D) than those situations when there were a lot of small events and one large event (which tended to generate a “ridge”, see Doncaster TSS in Appendix D). It is therefore likely that, despite being correlated, a larger data set would produce a more defined solution for this analysis. This adds further strength to the argument that lack of extensive, quality data has proved to be one of the main hurdles in achieving success in urban stormwater modelling.

When input parameters are investigated it is no surprise that the random errors in rainfall collection have the greatest potential impact on model uncertainty. Because of the model structure (higher b term for TSS), errors are amplified for TSS when compared to TP or TN. The relative errors from the input parameters are similar in magnitude to those calculated for monitoring data for EMCs (approximately 30%).

8.5 Future perspectives

The method developed addresses the original goals of the study and provides significant new evidence to support the hypothesis presented in Chapter 2. The model is able to predict event loads to a degree within the uncertainty of the measured loads. For TSS and TP a predictive model based on the sum of the rainfall intensity over an event is the preferred approach. For TN the message is less clear, and an approach based on total rainfall gives satisfactory results. This can be explained by the proportion of pollutants in a particulate phase. Greater intensity, and therefore kinetic energy, is needed to remove particulate pollutants, dissolved fractions are influenced by the contact with the stormwater itself. The within event analysis confirms that the primary processes responsible for pollutant generation are being addressed, but raises the question of when short timestep prediction would be needed. Accurate prediction of lag time is a major hurdle and needs to be considered against the modelling objectives. Although the uncertainty around rainfall measurement is an obvious area for improvement, the model could be used almost immediately in the eastern Melbourne area. It also provides a clear data set for comparison when parameters are developed for other regions.

There are several gaps that remain. Despite the strength of the rainfall intensity/pollutant load relationships for the studied catchments, and the results from the secondary sources that suggest the relationship holds, the method should be tested in different types of urban areas and for different hydrological conditions. The results should also be expanded to different pollutants, particularly metals.

Further work is already underway on the model structure and the uncertainty of various model components. This work will greatly inform model end users on its capabilities for specific uses.

Lastly, further work into the ecological and health impacts of stormwater is fundamental to ensure that models such as the one presented here are well designed and targeted for management uses.

9 BIBLIOGRAPHY

Aalderink, R. H., E. H. S. Van Duin,Peels, C.E., M.J.M. Scholten, (1990). Some Characteristics of Run-off Quality from a Separated Sewer System in Lelystad, The Netherlands. Proceedings of the Fifth International Conference on Urban Storm Drainage, Osaka, Japan.

Ahyerre, M., G. Chebbo, B.Tassin, , E.G aume, , (1998). "Stormwater quality modelling, an ambitious objective?" Water Science Technology **37**(1): 205-213.

Akan, A. O. (1987). "Pollutant Washoff by Overland Flow." Journal of Environmental Engineering **113**(4): 811-823.

Alley, W. M., F. W. Ellis, R.C. Sutherland, (1980). Toward a More Deterministic Urban Runoff-Quality Model. Proceedings of the International Symposium on Urban Storm Runoff, Lexington, Kentucky.

Alley, W. M. and P. E. Smith (1981). "Estimation of Accumulation Parameters for Urban Runoff Quality Modelling." Water Resources Research **17**(6): 1657-1664.

Allison, R. A., F. H. S. Chiew, T.A.McMahon,. (1997). Stormwater Gross Pollutants. Melbourne, CRC Catchment Hydrology.

Arnberg-Nielson, K. and P. Harremoes (1996). "The importance of inherent uncertainties in state-of-the-art urban storm drainage modelling for ungauged small catchments." Journal of Hydrology **179**: 305-319.

Asplund, R., B. W. Mar. (1982). "Total Suspended Solids in Highway Runoff in Washington State." Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers **108**(EE2): 391-404.

Baffaut, C. and J. W. Delleur (1990). "Calibration of SWMM Runoff Quality Model with Expert System." Journal of Water Resources Planning and Management **116**(2): 247-261.

Ball, J. E., R. Jenks, R., D.Auburg (1998). "An assessment of the availability of pollutant constituents on road surfaces." The Science of the Total Environment **209**: 243-254.

Bannerman, R. T., D. W. Owens, R.B. Dodds, N.J. Hornewer (1993). "Sources of Pollutants in Wisconsin Stormwater." Water Science and Technology **28**(3-5): 241-259.

Barfield, B. J., T. Y. Kao, C.T. Haan (1978). Sediment Control from Urban Construction Areas in the U.S. Urban Storm Drainage, Proceedings of the International Conference held at the University of Southampton, Pentech Press.

Beck, M. B. (1987). "Water quality modeling: A review of the analysis of uncertainty." Water Resources Research **23**(8): 1393-1442.

Bedient, P. B., J. L. Lambert, N.K. Springer., (1980). "Stormwater Pollutant Load-Runoff Relationships." Journal of the Water Pollution Control Federation **52**(9): 2396-2404.

Berndtsson, R. and J. Niemczynowicz (1988). "Spatial and temporal scales in rainfall analysis - some aspects and future perspectives " Journal of Hydrology **100**: 293-313.

Bertrand-Krajewski, J. L. and J. P. Bardin (2002b). Uncertainties and representativity of measurements in stormwater storage tanks. 9th ICUD, Portland, Portland, Oregon, USA.

Bertrand-Krajewski, J. L., S. Barraud, et al. (2002). "Uncertainties, performance indicators and decision aid applied to stormwater facilities." Urban Water **4**: 163-179.

Bertrand-Krajewski, J. L., G. Chebbo, A. Saget., (1998). "Distribution of pollutant mass vs volume in stormwater discharges and the first flush phenomenon." Water Research **32**(8): 2341-2356.

Bertrand-Krajewski, J. L. and O. Scrivener (1993). "Sewer sediment production and transport modelling: A literature review." Journal of Hydraulic Research **31**(4): 435-460.

Bertrand-Krajewski, J. L., O. Scrivener, et al. (1993). "Sewer sediment production and transport modelling: A literature review." Journal of Hydraulic Research **31**(4): 435-460.

Boyd, M. J., M. C. Bufill, et al. (1994). "Pervious and impervious runoff in urban catchments." Hydrological Sciences Journal **38**(6): 463-478.

Bradford, W. L. (1977). "Urban Stormwater Pollutant Loadings: A Statistical Summary Through 1972." Journal of the Water Pollution Control Federation **49**(4): 613-622.

Brezonik, L. and T. H. Stadelmann (2002). "Analysis and predictive models of stormwater runoff volumes, loads, and pollutant concentrations from watersheds in the Twin Cities metropolition area, Minnesota, USA." Water Research **36**: 1743-1757.

Brodie, I. (2007). "Prediction of stormwater particle loads from impervious urban surfaces based on a rainfall detachment index." Water Science and Technology **55**(4): 49-55.

Butler, D., S. Thedchanamoorthy, J.A Payne. (1992). "Aspects of Surface Sediment Characteristics on an Urban Catchment in London." Water Science and Technology **25**(8): 13-19.

Butts, M. B., Payne, J.T., Kristensen, M., Madsen, H., (2004). "An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation." Journal of Hydrology **298**: 242-266.

Calabro, P. S. (2001). "Cosmoss:conceptual simplified model for sewer system simulation A new model for urban runoff quality." Urban Water **3**: 33-42.

Castaldi, F. J. (1983). Final Report of the Nationwide Urban Runoff Program in Austin, Texas. Austin, USEPA.

Characklis, W. G., P. B. Bedient, F.G.Gaudet, F.L. Roe., (1978). "A Study of Stormwater Runoff Quality." Progress in Water Technology **10**(5/6): 673-699.

Charbeneau, R. J. and M. E. Barrett (1998). "Evaluation of Methods for Estimating Stormwater Pollutant Loads." Water Environment Research **70**(7): 1295-1302.

Chaubey, I., C. T. Haan, C.T. Grunwald, J.M. Salisbury (1999). "Uncertainty in the model parameters due to spatial variability of rainfall." Journal of Hydrology **220**: 48-61.

Chiew, F. H. S., H. P. Duncan, W. Smith (1997). Modelling Pollutant Buildup and Washoff: Keep It Simple. Wai-Whenua 1997, 24th Hydrology and Water Resources Symposium, Auckland, New Zealand, New Zealand Hydrological Society.

Chiew, F. H. S. & McMahon, T. A. (1997b). Modelling Daily Runoff and Pollutant Load from Urban Catchments. *Water (AWWA Journal),* 24**,** 16-17.

Chiew, F. H. S. and T. A. McMahon (1998). Estimation of Stormwater Runoff and Diffuse Pollution Loads. HydraStorm 98, 3rd International Symposium on Stormwater Management, Adelaide, South Australia, IEAust.

Choi, K. and J. E. Ball (2002). "Parameter estimation for urban runoff modelling." Urban Water(4): 31-41.

Chui, T. W., B. W. Mar (1982). "Pollutant Loading Model for Highway Runoff." Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers **108**(EE6): 1193-1210.

City Design (2003). Stormwater Monitoring Program 2002-2003. Brisbane, City Design/Brisbane City Council.

Clark, S., Pitt, R., Burian, S., Field, S., Fan, E., Heaney, J., Wright, L. (2006). Annotated bibliography of wet weather flow literature from1996 through 2004. Middleton, PA.

Coleman, T. J. (1993). A Comparison of the Modelling of Suspended Solids Using SWMM3 Quality Prediction Algorithms with a Model Based on Sediment Transport Theory. Proceedings of the Sixth International Conference on Urban Storm Drainage, Niagara Falls, Ontario, Canada.

Cordery, I. (1977). "Quality Characteristics of Urban Stormwater in Sydney, Australia." Water Resources Research **13**(1): 197-201.

Cowen, W. F. and G. F. Lee (1976). "Phosphorus Availability in Particulate Materials Transported by Urban Runoff." Journal of the Water Pollution Control Federation **48**(3): 580-591.

CRC for Freshwater Ecology (2004). Urban stormwater and the ecology of streams. Canberra, Cooperative Research Centre for Freshwater Ecology.

CRCCH (2003). MUSIC users guide: Version 2, Co-operative Research Centre for Catchment Hydrology.

Cristina, C. M. and J. J. Sansalone (2003). ""First Flush", power law and particle seperation diagrams for urban storwater suspended particulates." Journal of Environmental Engineering **April 2003**: 298-307.

Cristina, C. M. and J. J. Sansalone (2003). "‘‘First Flush,’’ Power Law and Particle Separation Diagrams for Urban Storm-Water Suspended Particulates." Journal of Environmental Engineering **April 2003**: 298-307.

CSIRO (1996). Port Phillip Bay Environmental Study, CSIRO.

Deletic, A., C. Maksimovic, M. Ivetic (1997). "Modelling of Storm Washoff of Suspended Solids from Impervious Surfaces." Journal of Hydraulic Research **35**(1): 99-118.

Deletic, A. and D. W. Orr (2003). Water quality of runoff from paved surfaces. IAHR Congress, Thessaloniki, Greece.

Deletic, A. B. (1998). "Evaluation of Water Quality Factors in Storm Runoff from Paved Areas." Journal of Environmental Engineering **124**(9): 869-879.

Deletic, A. B. (1998). "The first flush load of urban surface runoff." Water Research **32**(8): 2462-2470.

Desbordes, M. and E. Servat (1984). Solids in Urban Runoff - Statistical Analysis of French Experimental Data. Proceedings of the Third International Conference on Urban Storm Drainage, Goteborg, Sweden.

Desbordes, M. and E. Servat (1984). Solids in Urban Runoff: Statistical Analysis of French Experimental Data. 3rd International Conference on USD. 1984 Goteborg, Goteborg.

Desbordes, M. and E. Servat (1987). Transport Model over Urban Catchment Surfaces Regarding TSS, BOD5, and COD. Urban Storm Water Quality, Planning and Management, Proceedings of the Fourth International Conference on Urban Storm Drainage, Lausanne, Switzerland.

Deutsch, J. C. and M. Desbordes (1981). Study of Runoff Pollution for Urban Planning. Urban Stormwater Quality, Management, and Planning, Proceedings of the Second International Conference on Urban Storm Drainage, Urbana, Illinois, Water Resources Publications, Colorado.

Deutsh, J. C. and J. C. Hemain (1984). Main results of the French National Programme of Urban Runoff Quality Measurement. 3rd International Conference on USD. 1984 Goteborg, Goteborg.

Dotto, C., Deletic, A., Fletcher, T.D., (2009). "Analysis of parameter uncertainty of a flow and quality stormwater model." Water Science and Technology **60**(3): 717-725.

Driscoll, E. D. (1986). Lognormality of Point and Non-Point Source Pollutant Concentrations. Urban Runoff Quality - Impact and Quality Enhancement Technology, Proceedings of an Engineering Foundation Conference, New Hampshire, ASCE.

Driver, N. E. (1990). Summary of Nationwide Analyses of Storm-Runoff Quality and Quantity in Urban Watersheds. Proceedings of the Fifth International Conference on Urban Storm Drainage, Osaka, Japan.

Driver, N. E. and B. M. Troutman (1989). "Regression Models for Estimating Urban Storm-Runoff Quality and Quantity in the United States." Journal of Hydrology **109**: 221-236.

Duchon, C. E. and G. R. Essenberg (2001). "Comparative rainfall observations from pit and aboveground rain gauges with and without wind shields." Water Resources Research **37**(12): 3253-3263.

Duda, A. M., D. R. Lenat, et al. (1982). "Water Quality in Urban Streams - What We Can Expect." Journal of the Water Pollution Control Federation **54**(7): 1139-1147.

Duncan, H. P. (1995). A Review of Urban Storm Water Quality Processes. Melbourne, Australia, Cooperative Research Centre for Catchment Hydrology.

Duncan, H. P. (1995a). A Bibliography of Urban Storm Water Quality. Melbourne, Australia, Cooperative Research Centre for Catchment Hydrology.

Duncan, H. P. (1999). Urban Stormwater Quality: A Statistical Overview. Melbourne, Australia, Cooperative Research Centre for Catchment Hydrology.

Ellis, J. B. (1986). Pollutional aspects of urban runoff. Urban Runoff Pollution. H. C. Torno, J. Marsalek and M. Desbordes. Berlin, Springer-Verlag. **G10:** 1-38.

Engineers Australia (1997). Australian Rainfall and Runoff Crows Nest Engineers Australia.

Fergusson, J. E. and D. E. Ryan (1984). "The Elemental Composition of Street Dust from Large and Small Urban Areas Related to City Type, Source and Particle Size." The Science of the Total Environment **34**: 101-116.

Fowler, J., L. Cohen, et al. (1998). Practical Statistics for Field Biology. Chichester, England, John Wiley and Sons.

Francey, M., A. Deletic, T. Fletcher, H.P. Duncan (2004). Monitoring of Stormwater Runoff in Melbourne. 19th European Junior Scientists Workshop - Process Data and Integrated Urban Water Modelling, Lyon, France.

Freund, A. P. and C. D. Johnson (1980). Comparison and Relationships of Stormwater Quality and Basin Characteristics: Madison, Wisconsin. Proceedings of the International Symposium on Urban Storm Runoff, Lexington, Kentucky.

Fuchs, S., H. Brombach, et al. (2004). New database on urban runoff pollution. Novatech, Lyon, France.

Gaume, E., J. Villeneuve, M. Desbordes (1998). "Uncertainty assessment and analysis of the calibrated parameter values of an urban storm water quality model." Journal of Hydrology **210**: 38-50.

Geiger, W. (1987). Flushing effects in combined sewer systems. 4th International Conference on urban Drainage, Lausanne, Switzerland.

Geiger, W. F. (1984). Goodness and Limits in Urban Runoff Modelling. Proceedings of the Third International Conference on Urban Storm Drainage, Goteborg, Sweden.

GeoSyntec Consultants and ASCE (2002). Urban Stormwater BMP Performance Monitoring. Washington DC, US EPA.

GHD and EPA Vic (1981). Characterisation of Pollution in Urban Stormwater Runoff. Canberra, Australian Water Resources Council.

Ginzburg, L. R. and C. X. J. Jensen (2004). "Rules of thumb for judging ecological theories." Trends in Ecology and Evolution **19**(3): 121-126.

Grayson, R., I. D. Moore, et al. (1992). "Physically based hydrologic modelling 2. Is the concept realistic?" Water Resources Research **26**(10): 2659-2666.

Grottker, M. (1987). "Runoff Quality from a Street with Medium Traffic Loading." The Science of the Total Environment **59**: 457-466.

Gutteridge Haskins and Davey and Environment Protection Authority of Victoria (1981). Characterisation of Pollution in Urban Stormwater Runoff. Canberra, Australian Water Resources Council.

Hammer, T. R. (1972). "Stream Enlargement due to Urbanisation." Water Resources Research **8**(6): 1530-1540.

Harmel, R. D., Cooper, R.J., Slade, R.M., Haney, R.L., Arnold, J.G., (2006). "Cumulative uncertainty in measured streamflow and water quality data for small watersheds." Transactions of the American Society of Agricultural and Biological Engineers **43**(3): 689-701.

Harmel, R. D., King, K. W., Wolfe,J. E., Torbert, H. A (2002). "Minimum flow considerations for automated storm sampling on small watersheds." Texas Journal of science **54**(2): 177-188.

Harrison, R. M. (2002). "Key Pollutants - Airborne Particles." Clean Air and Environmental Protection **32**(4): 38-41.

Hartigan, J. P., A. M. Lumb, J.T. Smullen, T.J.Grizzard, P.A. Alcivar (1978). Calibration of Urban Nonpoint Pollution Loading Models. Verification of Mathematical and Physical Models in Hydraulic Engineering, Proceedings of the 26th Annual Hydraulics Division Specialty Conference, University of Maryland, Hydraulics Division, ASCE, New York.

Haster, T. W. and W. P. James (1994). "Predicting Sediment Yield in Storm-Water Runoff from Urban Areas." Journal of Water Resources Planning and Management **120**(5): 630-650.

Hatt, B. E., T. D. Fletcher, C.J. Walsh, S.L. Taylor (2004). "The Influence of Urban Density and Drainage Infrastructure on the Concentrations and Loads of Pollutants in Small Streams." Environmental Management **34**(1): 112-124.

House, M. A., Ellis, J. B., Herricks, E. E., Hvitved-Jacobsen, T., Seager, J., Lijklema, L., Aalderink, H. and Clifforde, I. T (1993). "Urban Drainage - Impacts on Receiving Water Quality." Water Science and Technology **27**(12): 117-158.

Huber, W. C. (1986). Modelling Urban Runoff Quality: State of the Art. Urban Runoff Quality - Impact and Quality Enhancement Technology, Proceedings of an Engineering Foundation Conference, New Hampshire, ASCE, New York.

Huber, W. C. (1992). Prediction of Urban Nonpoint Source Water Quality: Methods and Models. International Symposium on Urban Stormwater Management, Sydney.

Huber, W. C. and R. E. Dickinson (1988). Storm Water Management Model, Version 4: Users' Manual. Gainesville, University of Florida.

Huber, W. C. and J. P. Heaney (1980). Operational Models for Stormwater Quality Management. Environmental Impact of Nonpoint Source Pollution. M. J. Overcash and J. M. Davidson, Ann Arbor Science Publishers Inc**:** 397-444.

Hudson, N. (1971). Soil Conservation. London, B T Batsford Ltd.

Hvitved-Jacobsen, T., Yousef, Y. A., Wanielista, M. P. and Pearce, D. B. (1984). "Fate of Phosphorus and Nitrogen in Ponds Receiving Highway Runoff." The Science of the Total Environment **33**: 259-270.

Irish, L. B., Barrett, M. E., Malina, J. F. and Charbeneau, R. J. (1998). "Use of Regression Models for Analyzing Highway Stormwater Loads." Journal of Environmental Engineering(October 1998): 987-993.

Jakeman, A. J. and G. M. Hornberger (1993). "How Much Complexity is Warranted in a Rainfall-Runoff Model." Water Resources Research **29**(8): 2637-2649.

Jakeman, A. J., Taylor, J. A. and Symons, H. D (1992). "Efficient Tools for Analysising the Influence of Sources and Meteorology on Urban Ambient COncentration Trends Illustrated for Canberra, Australlia." Ecological Modelling **64**: 125-157.

James, W. and B. Shivalingaiah (1986). Continuous Mass Balance of Pollutant Build-up Processes. Urban Runoff Pollution. H. C. Torno, J. Marsalek and M. Desbordes. Berlin, Springer-Verlag. **G10:** 243-271.

Jewell, T. K. and D. D. Adrian (1982). "Statistical Analysis to Derive Improved Stormwater Quality Models." Journal of the Water Pollution Control Federation **54**(5): 489-499.

Jones, J. E. (1986). Urban Runoff Impacts on Receiving Waters. Urban Runoff Quality - Impact and Quality Enhancement Technology, Proceedings of an Engineering Foundation Conference, New Hampshire, ASCE.

Kanso, A., M. C. Gromaire, et al. (2003). "Bayesian approach for the calibration of models: Application to an urban stormwater polltion model." Water Science and Technology **47**(4): 77-84.

Kanso, A., B. Tassin, et al. (2004). Managing uncertainties in urban runoff quality models: A benching methodolgy. Novatech, Lyon, France.

Kanso, A., B. Tassin, et al. (2004). Managing uncertainties in urban runoff quality models: A benchmarking methodolgy. Novatech. Lyon, France.

Kenney, B. C. (1982). "Beware of spurious self-correlations." Water Resources Research **18**(4): 1041-1048.

Kuczera, G., Kavetski, D., Franks, S., Thyer, M., (2006). "Towards a bayesian total error analysis of conceptual rainfall-runoff models: characterising model error using storm-dependent parameters." Journal of Hydrology **331**(1-2): 161-177.

Kuo, J. T., Yan, Y. L. and Li, K. C. (1993). "A Simplified Computer Model for Nonpoint Source Pollution in a Small Urban Area." Water Science and Technology **28**(3-5): 701-706.

LaBarbera, P., L. G. Lanza, et al. (2002). "Tipping bucket mechanical errors and their influence on rainfall statistics and extremes." Water Science and Technology **45**(2): 1-9.

Lee, D. S., J. A. Garland, et al. (1994). "Atmospheric Concentrations of Trace Elements in Urban Areas of the Unitied Kingdom." Atmos Environ **28**(16): 2691-2713.

Lee, J. H., Bang, K. W., Ketchum, L. H., Choe, M. J. and Yu, M. J. (2002). "First flush analysis of urban storm runoff." The Science of the Total Environment(293): 163-175.

Leecaster, M. K., Schiff, K. and Tienfenthaler, L. L. (2002). "Assessment of efficient sampling designs for urban stormwater monitoring." Water Research **36**: 1556-1564.

Legates, D. R. and G. J. McCabe (1999). "Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation." Water Resources Research **35**(1): 233-241.

Lloyd, S. D. (2004). Exploring the opportunities and barriers to sustainable stormwater management practices in residential catchments. Department of Engineering. Melbourne, Monash.

Maksimovic, C. and Radojkovic (1986). Urban Drainage Catchments: Selected worldwide rainfall-runoff data from experimental catchments., Pergamon Press, Belgrade.

Malmquist, P. (1978). "Atmospheric Fallout and Street Cleaning - Effects on Urban Storm Water and Snow." Progress in Water Technology **10**(5/6): 495-505.

Mance, G. and M. M. I. Harman (1978). The Quality of Urban Storm-Water Runoff. Urban Storm Drainage, Proceedings of the International Conference held at the University of Southampton, Southampton, Pentech Press, London.

MapInfo Corporation (2002). MapInfo Professional Version 7.0.

Marsalek, J. (1981). "Calibration of the tipping-bucket raingauge." Journal of Hydrology **53**: 343-354.

Mason, Y., Ammann, A. A., Ulrich, A. and Sigg, L. (1999). "Behaviour of heavy metals, nutrients, and major components during roof runoff infiltration." Environmental Science Technology **33**: 1588-1597.

McCarthy, D. T., Deletic, A., Mitchell, V.G., Fletcher, T.D., Diaper, C., (2008). "Uncertaities in stormwater E. coli levels." Water Research **42**(6-7): 1812-1824.

McPherson, T. N., S. J. Burian, et al. (2002). "Comparison of the pollutant loads in dry and wet weather runoff in a southern California urban watershed." Water Science and Technology **45**(9): 225-261.

McPherson, T. N., Burian, S. J., Turin, H.J., Strenstron, M.K., Suffet, I.H., (2002). "Comparison of the pollutant loads in dry and wet weather runoff in a southern California urban watershed " Water Science and Technology **45**(9): 255-261.

Melways Publishing (2002). Melways Greater Melbourne Street directory. Melbourne, Melways Publishing Pty. Ltd.

Meyer, L. D., Wischmeier, W.H., (1969). "Mathematical simulation of the process of soil erosion by water." Trans. Amer. Soc. Agric. Eng **12**: 756-762.

Miller, R. A. and H. C. Mattraw (1982). "Storm Water Runoff Quality from Three Land-Use Areas in South Florida." Water Resources Bulletin **18**(3): 513-519.

Molini, A., Lanza, L. G. and LaBarbera, P. "The impact of tipping-bucket raingauge measurement errors on design rainfall for urban-scale applications." Hydrological Processes **19**: 1073-1088.

Moreira, L. F. F., Righetto, A. M. and Medeiros, V. M. (2006). Uncertainty analysis associated with Rainfall Spatial Distribution in an Experimental Semiarid Watershed, Northeastern Brazil. 3rd Biennial meeting of the International Environmental Modelling and Software Society. Vermont, USA.

Muthukaruppan, M., F. H. S. Chiew, et al. (2002). Size Distribution and Partitioning of Urban Pollutants. 9th International Conference on Urban Drainage, Portland, Oregan, USA, ASCE.

Nakamura, E. (1984). Factors Affecting Stormwater Quality Decay Coefficient. Proceedings of the Third International Conference on Urban Storm Drainage, Goteborg, Sweden.

Nash, J. E. and J. V. Sutcliffe (1970). "River forecasting through conceptual models, 1. A discussion of principles." Journal of Hydrology **10**: 282-290.

Nelen, F., A. Mooijman, et al. (1992). "The importance of rainfall distribution in urban rainfall operation." Nordic Hydrology **23**: 121-136.

North Shore City Council (2002). Project CARE: 2001 Monitoring. Auckland, North Shore City Council.

Novotny, V., Sung, H.-M., Bannerman, R. and Baum, K. (1985). "Estimating Nonpoint Pollution from Small Urban Watersheds." Journal of the Water Pollution Control Federation **57**(4): 339-348.

Odnevall-Wallinder, W. H. I., Leygraf, C., (2000). "A laboratory study of copper and zinc runoff during first flush and steady state conditions." Corrosion Science **43**(1): 127-146.

Ogden, F. L. and P. Julien, Y. (1993). "Runoff Sensitivity to Temporal and Spacial Rainfall Variability at Runoff Plane and Small Basin Scales." Water Resources Research **29**(8): 2589-2597.

Orr, D. W. (2002). Pollution Release from Urban Asphalt Surfaces, Aberdeen University.

Orr, D. W. and A. B. Deletic (2003). Pollution Build up on road surfaces. Aberdeen.

Pitt, R. (1979). Demonstration of Nonpoint Pollution Abatement Through Improved Street Cleaning Practices. Cincinnati, United States Environmental Protection Agency.

Pitt, R. and A. Maestre (2005). Stormwater quality as described in the National Stormwater Quality Database (NSQD). 10th ICUD, Copenhagen, Denmark.

Prairie, Y. T. and D. F. Bird (1989). "Some misconceptions about the spurious correlation problem in the ecological literature." Oecologia **81**: 285-288.

Price, R. K. and G. Mance (1978). A Suspended Solids Model for Storm Water Runoff. Urban Storm Drainage, Proceedings of the International Conference held at the University of Southampton, Pentech Press.

Randall, C. W., Ellis, K., Grizzard, T. J. and Knocke, W. R. (1982). Urban Runoff Pollutant Removal by Sedimentation. Stormwater Detention Facilities - Planning, Design, Operation, and Maintenance, Henniker, New Hampshire, ASCE.

Randall, C. W., T. J. Grizzard, et al. (1981). Comparison of Pollutant Mass Loads in Precipitation and Runoff in Urban Areas. Urban Stormwater Quality, Management, and Planning, Proceedings of the Second International Conference on Urban Storm Drainage, Urbana, Illinois, Water Resources Publications, Colorado.

Reinertsen, T. R. (1981). Quality of Stormwater Runoff from Streets. Urban Stormwater Quality, Management, and Planning, Proceedings of the Second International Conference on Urban Storm Drainage, Urbana, Illinois, Water Resources Publications, Colorado.

RossRakesh, S., C. Gippel, et al. (1999). Blackburn Lake discharge and water quality monitoring program:Data summary and interpretation. Melbourne, CRC for Catchment Hydrology/Freshwater Ecology.

Salles, C., Poesen, J. and Govers, G. (2000). "Statistical and Physical Analysis of Soil Detachment by Raindrop Impact: Rain Erosivity Indices and Threshold Energy." Water Resources Research **36**(9): 2721-2729.

Sartor, J. D. and G. B. Boyd (1972). Water Pollution Aspects of Street Surface Contaminants. Washington, D.C., United States Environmental Protection Agency.

Salles, C., Poesen, J. and Govers, G. (1974). "Water Pollution Aspects of Street Surface Contaminants." Journal of the Water Pollution Control Federation **46**(3): 458-467.

Shaheen, D. G. (1975). Contributions of Urban Roadway Usage to Water Pollution. Washington, D.C., United States Environmental Protection Agency.

Shapiro, S. S. and M. B. Wilk (1965). "An analysis of variance test for normality." Biometrika **52**(3): 591-9.

Sharpin, M. G. (1993). Hydrological Variability, Urbanisation and Water Quality in Australia. Proceedings of the Sixth International Conference on Urban Storm Drainage, Niagara Falls, Ontario, Canada.

Sharpin, M. G. (1995). Stormwater Quality Characteristics from Urban and Non-Urban Catchments in South-Eastern Australia. AWWA 16th Federal Convention, Darling Harbour, Sydney, Australia, Australian Water and Wastewater Association.

Shivalingaiah, B. and W. James (1984a). Applied Algorithms for Pollutant Buildup, Washoff and Transport in Stormwater. Water Quality Evolution Within the Hydrological Cycle of Watersheds, Proceedings of the Canadian Hydrology Symposium No 15, Quebec, Canada.

Shivalingaiah, B. and W. James (1984b). Algorithms for Buildup, Washoff, and Routing Pollutants in Urban Runoff. Proceedings of the Third International Conference on Urban Storm Drainage, Goteborg, Sweden.

Shuster, W. D., J. Bonta, et al. (2005). "Impacts of impervious surface on watershed hydrology: A review." Urban Water Journal **2**(4): 263-275.

Smith, W. (1997). Estimation of pollution accumulation for stormwater washoff modelling. Civil Engineering. Melbourne, Monash.

Smullen, J. T., A. L. Shallcross, et al. (1999). "Updating the U.S. nationwide urban runoff quality database." Water Science Technology **39**(12): 9-16.

Spangberg, A. and J. Niemczynowicz (1993). Measurements of Pollution Wash-off from an Asphalt Surface. Proceedings of the Sixth International Conference on Urban Storm Drainage, Niagara Falls, Ontario, Canada.

Sprent, J. I., Parsons, R. (2000). "Nitrogen fixation in legume and non-legume trees

" Field Crops Research **65**(2-2).

Srinanthakumar, K. and G. P. Codner (1992). Factors Affecting Pollutant Washoff Decay Coefficient in Urban Catchments. International Symposium on Urban Stormwater Management, Sydney, Australia.

Srinanthakumar, K. and G. P. Codner (1992a). Data Evaluation Before Urban Runoff Quality Modelling. International Symposium on Urban Stormwater Management, Sydney.

Srinanthakumar, S. K. and G. P. Codner (1993). A Process Based Model for Pollution Washoff by Overland Flow. Hydrology and Water Resources Symposium, Newcastle, IEAust.

Strangeways, I. (2004). "Improving precipitation measurement." International Journal of Climatology **24**: 1443-1460.

Talylor, B. N., Kuyatt, C.E., (1994). Guidelines for evaluating and expressing the uncertainty of NIST measurement results. Technical Note 1296, National Institute of Standards and Technology, USA.

Tasker, G. D. and N. E. Driver (1988). "Nationwide Regression Models for Predicting Urban Runoff Water Quality at Unmetered Sites." Water Resources Bulletin **24**(5): 1091-1101.

Taylor, G. D., Fletcher, T. D., Wong, T. H. F., Breen, P. F. and Duncan, H. P. (2005). "Nitrogen composition in urban runoff—implications for stormwater management." Water Research **39**: 1982-1989.

Troutman, B. M. (1983). "Runoff prediction erros and bias in parameter estimation nduced by spatial variability of rainfall on storm runoff." Water Resources Research **19**(3): 791-810.

Tsihrintzis, V. A. and R. Hamid (1998). "Runoff quality prediction from small urban catchments using SWMM." Hydrological Processes **12**: 311-329.

van Dijk, A. I. J. M., Bruijnzeel, L. A. and Rosewell, C. J. (2002). "Rainfall intensity - kinetic energy relationships:a critical literature appraisal." Journal of Hydrology **261**(2002): 1-23.

Var, F., Y. Narita, et al. (2000). "The Concentration, Trend and Seasonal Variation of Metals in the Atmosphere in 16 Japenese Cities shown by the Results of National Air Surveillance Network (NASN) from 1974 to 1996." Atmos Environ **34**: 2755-2770.

Vaze, J. and F. H. S. Chiew (2002). "Experimental study of pollutant accumulation on an urban road surface." Urban Water **4**(2002): 379-389.

Vaze, J. and F. H. S. Chiew (2003). "Comparative evaluation of urban storm water quality models." Water Resources Research **39**(10): 1280.

Vaze, J. and F. H. S. Chiew (2003b). "Study of pollutant washoff from small impervious experimental plots." Water Resources Research **39**(6): 1160.

Vorreiter, L. and C. Hickey (1994). Incidence of the First Flush Phenomenon in Catchments of the Sydney Region. Water Down Under 94, Adelaide, South Australia.

Walden, W. J. (1999). Non-point source pollutant export estimation from urban catchments. Faculty of Engineering. Brisbane, University of Queensland.

Walsh, C. J., Fletcher, T., Leonard and Ladston, A. (2004). Urban stormwater and the ecology of streams. Canberra, Cooperative Research Centre for Freshwater Ecology.

Walsh, C. J., Sharpe, A. K., Breen, P. F. and Sonneman, J. A. (2001). "Effects of Urbanization on Streams of the Melbourne Region, Victoria, Australia. I. Benthic Macroinvertebrate Communities." Freshwater Biology(46).

Wang, X., Tong, S., Verall, K., Gerber, R. and Wolff, R. (2006). "Air Pollution Trends in Brisbane, Australia, between 1980 and 2003." Clean Air and Environmental Quality **40**(1): 34-39.

Weibel, S. R., Anderson, R. J. and Woodward, R. L. (1964). "Urban Land Runoff as a Factor in Stream Pollution." Journal of the Water Pollution Control Federation **36**(7): 915-924.

Wong, T. H. F., Breen, P. F. and Lloyd, S. D. (2000). Water Sensitive Road Design options for improving stormwater quality of road runoff, Cooperative research centre for catchment hydrology.

Xanthopoulos, C. and H. H. Hahn (1993). Anthropogenic Pollutants Wash-off from Street Surfaces. Proceedings of the Sixth International Conference on Urban Storm Drainage, Niagara Falls, Ontario, Canada.

Yaziz, M. I., Gunting, H., Sapari, N. and Ghazali, A. W. (1989). "Variations in Rainwater Quality from Roof Catchments." Water Research **23**(6): 761-765.

Young, R. A. and J. L. Wiersma (1973). "The role of rainfall impact in soil detachment and transport." Water Resources Research **9**(6): 1629-1636.

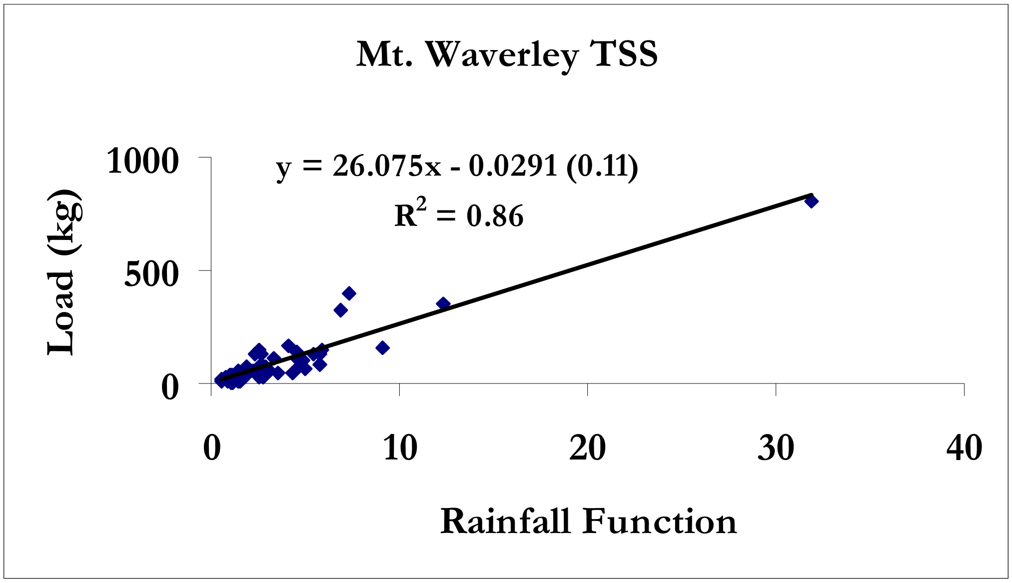
Yuan, Y., Hall, K. and Oldham, C. "A preliminary model for predicting heavy metal contaminent loading from an urban catchment." The Science of the Total Environment **266**: 299-307.

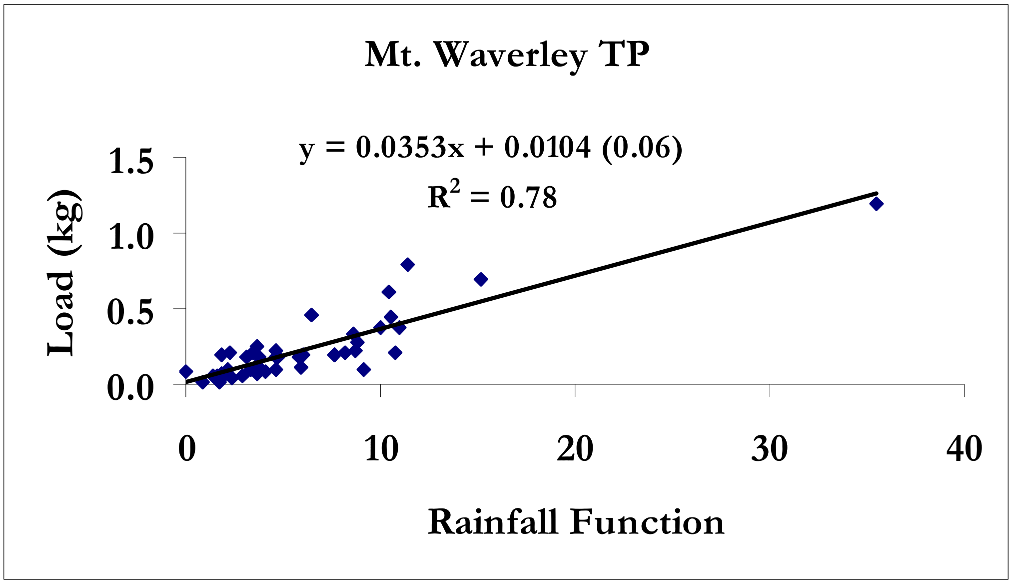
Zoppou, C. (2001). "Review of urban storm water models." Environmental Modelling & Software **16**: 195-231.

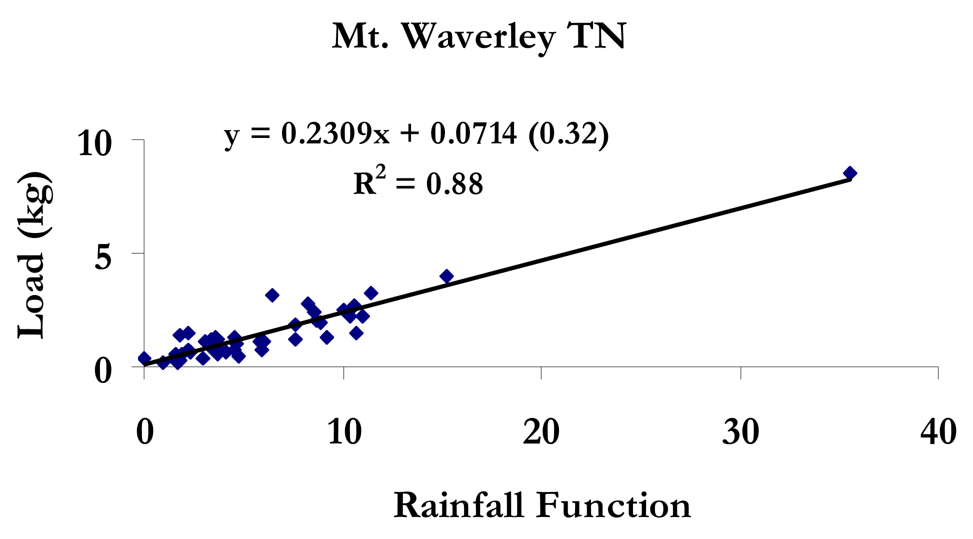
10 Appendix A - Significance of Y-intercepts

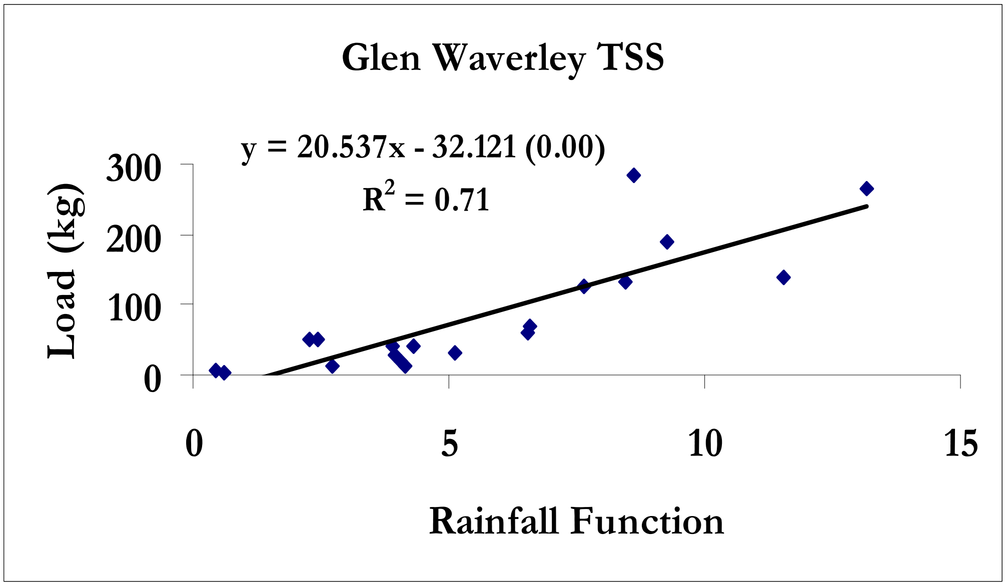
All plots show the Load versus the Rainfall function, with the equation of the line of best fit, the correlation and the significance of the y-intercept in brackets.

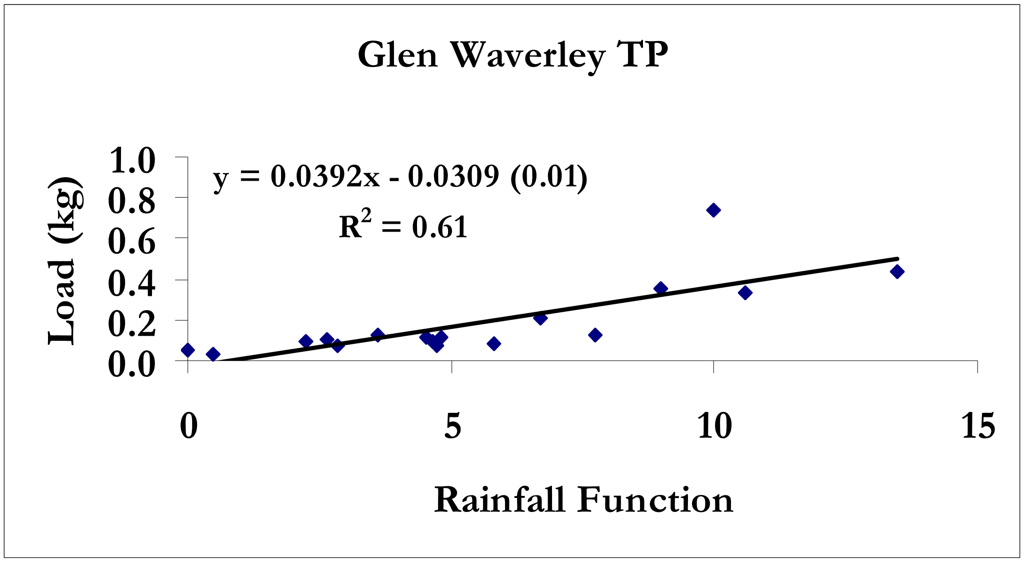
10.1 Primary data sets (TSS, TP, TN)

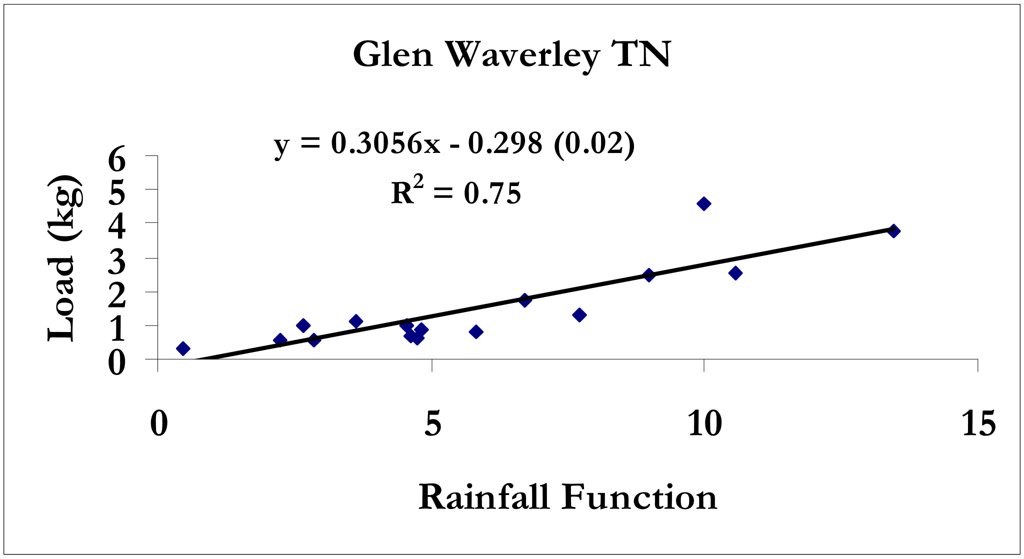


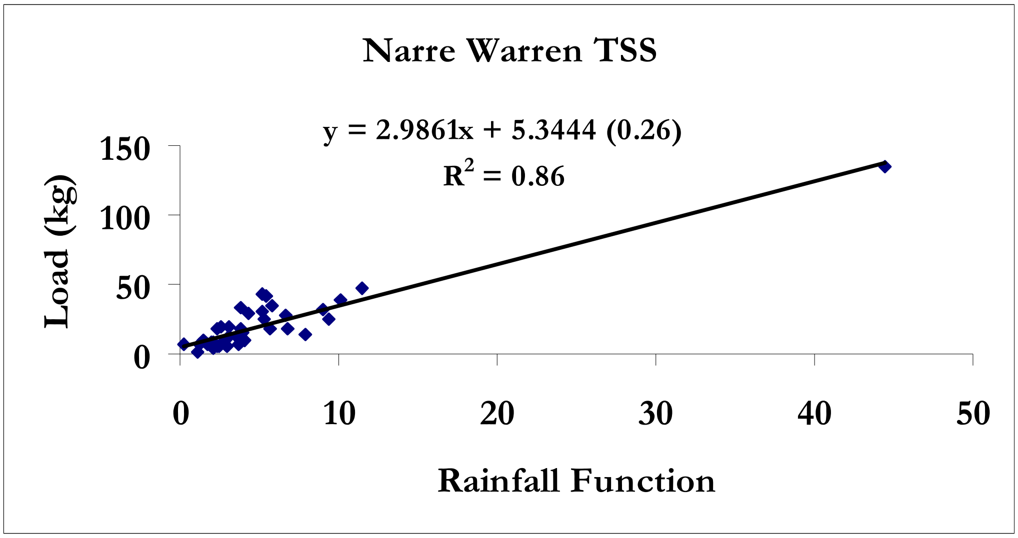


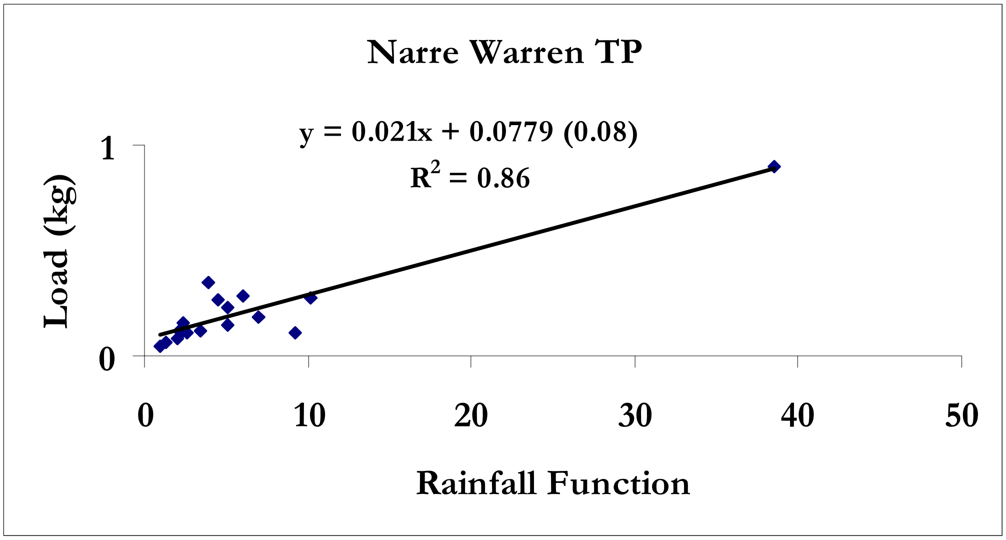

Mt. Waverley TN
y = 0.2309x + 0.0714 (0.32)
$R^2 = 0.88$
Load (kg)
10
5
0
0
10
20
30
40
Rainfall Function

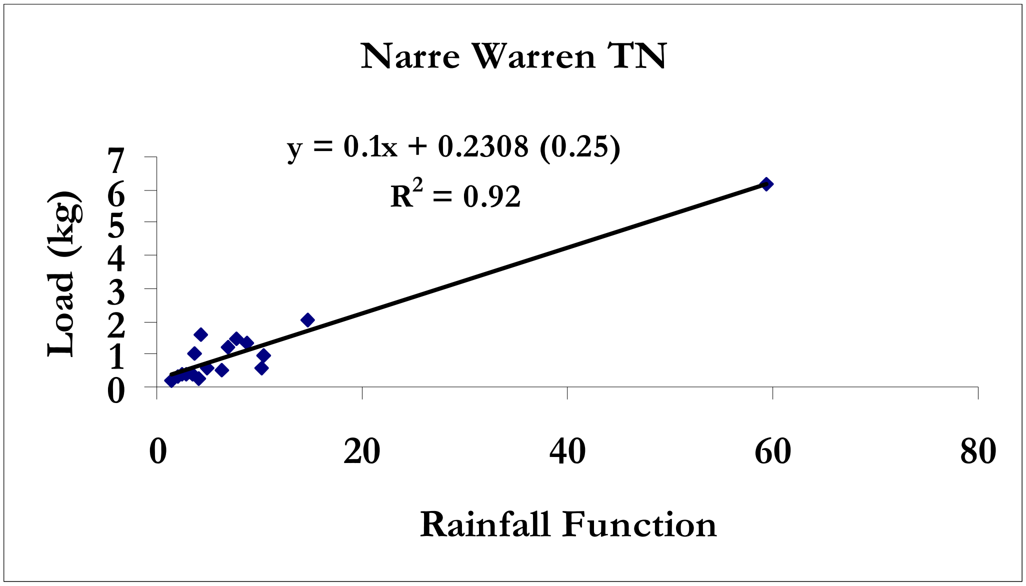

Glen Waverley TSS
y = 20.537x - 32.121 (0.00)
$R^2 = 0.71$
Load (kg)
300
200
100
0
0
5
10
15
Rainfall Function

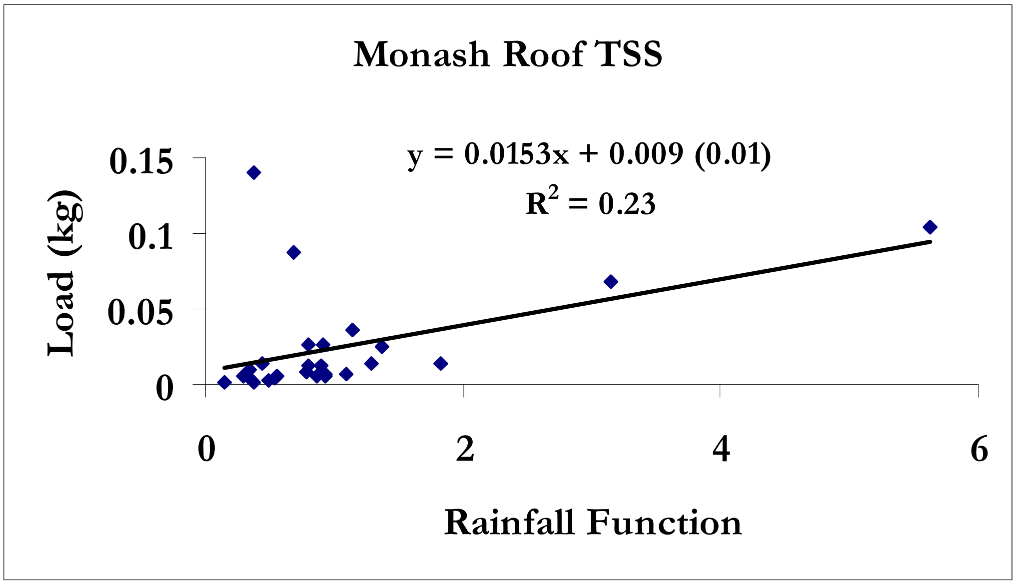

Glen Waverley TP
y = 0.0392x - 0.0309 (0.01)
R2 = 0.61
Load (kg)
1.0
0.8
0.6
0.4
0.2
0.0
0
5
10
15
Rainfall Function

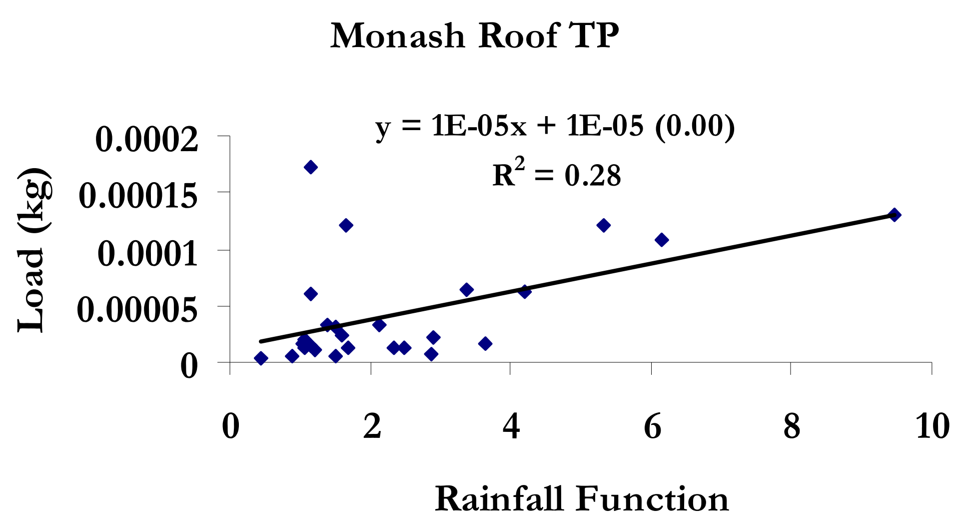

Glen Waverley TN
y = 0.3056x - 0.298 (0.02)
R2 = 0.75
Load (kg)
6
5
4
3
2
1
0
0
5
10
15
Rainfall Function

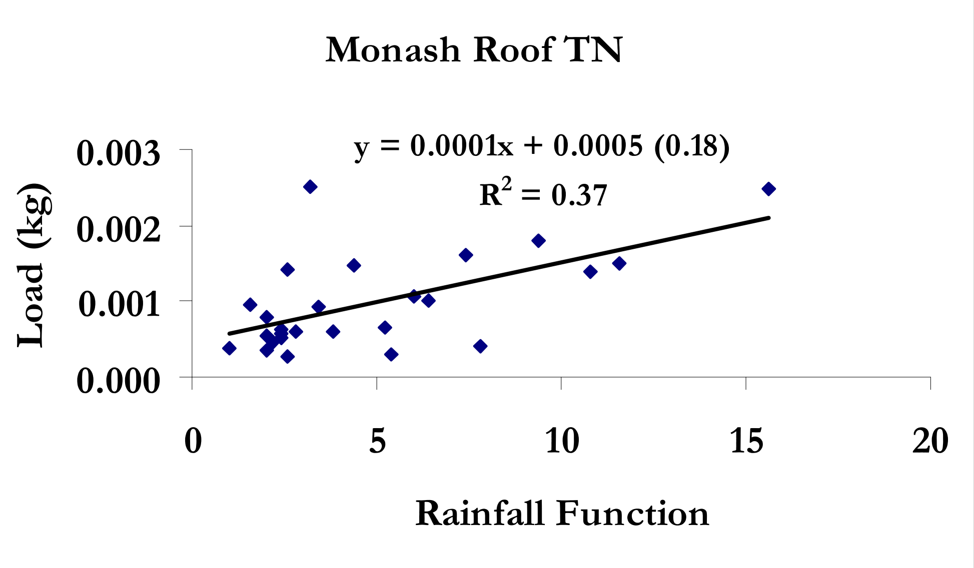

Narre Warren TSS
y = 2.9861x + 5.3444 (0.26)
R^2 = 0.86
Load (kg)
150
100
50
0
0
10
20
30
40
50
Rainfall Function

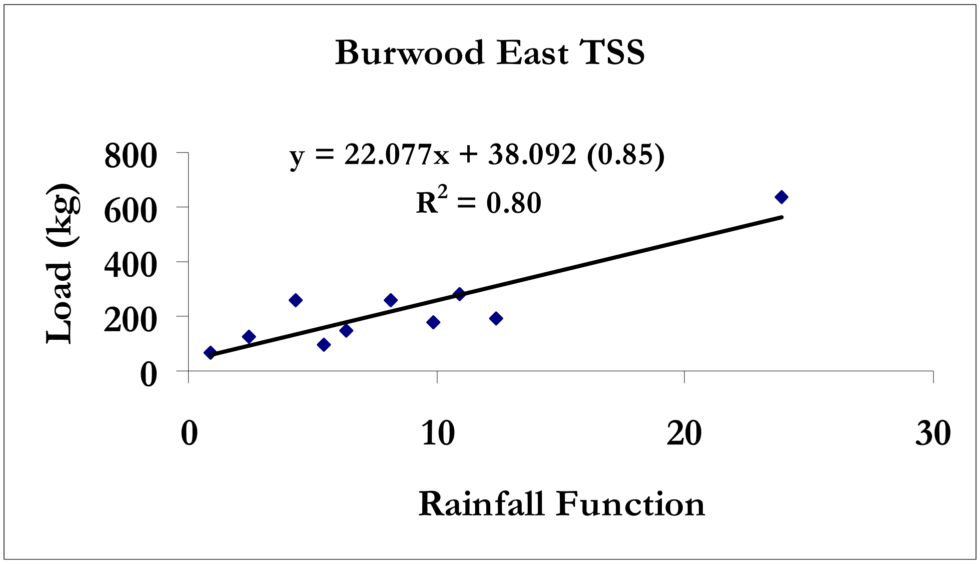

Narre Warren TP
y = 0.021x + 0.0779 (0.08)
R^2 = 0.86
Load (kg)
1
0
0
10
20
30
40
50
Rainfall Function

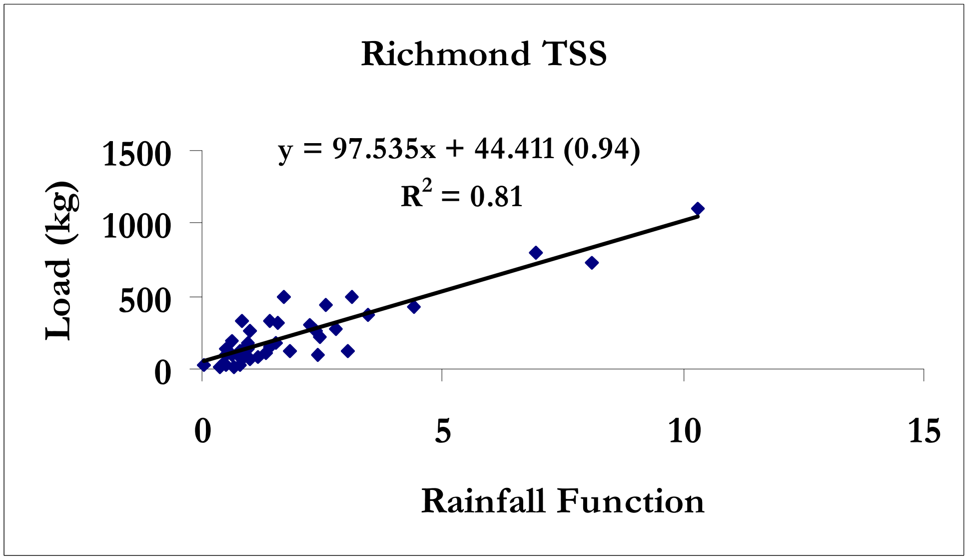

Narre Warren TN
y = 0.1x + 0.2308 (0.25)
$R^2 = 0.92$
Load (kg)
0
1
2
3
4
5
6
7
0
20
40
60
80
Rainfall Function

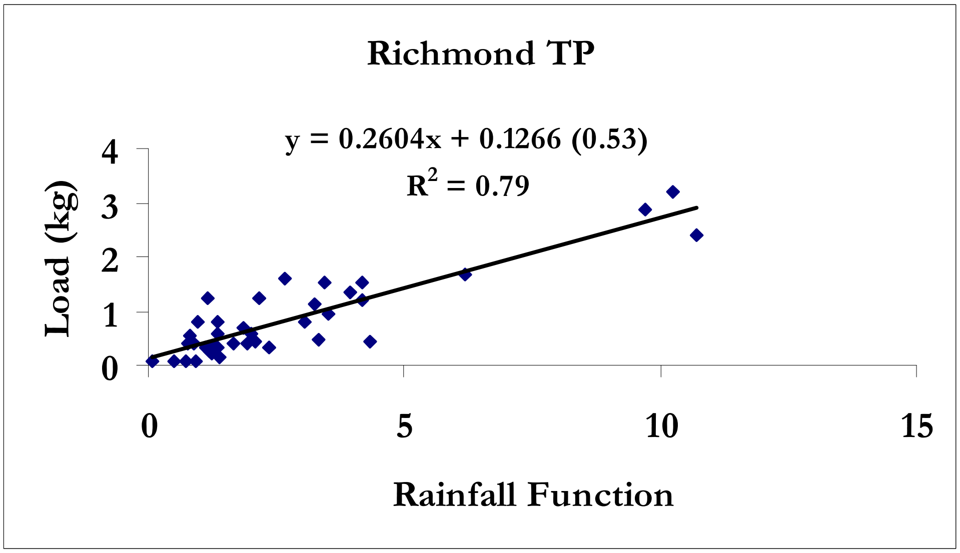

Monash Roof TSS
y = 0.0153x + 0.009 (0.01)
$R^2 = 0.23$
Load (kg)
0
0.05
0.1
0.15
0
2
4
6
Rainfall Function

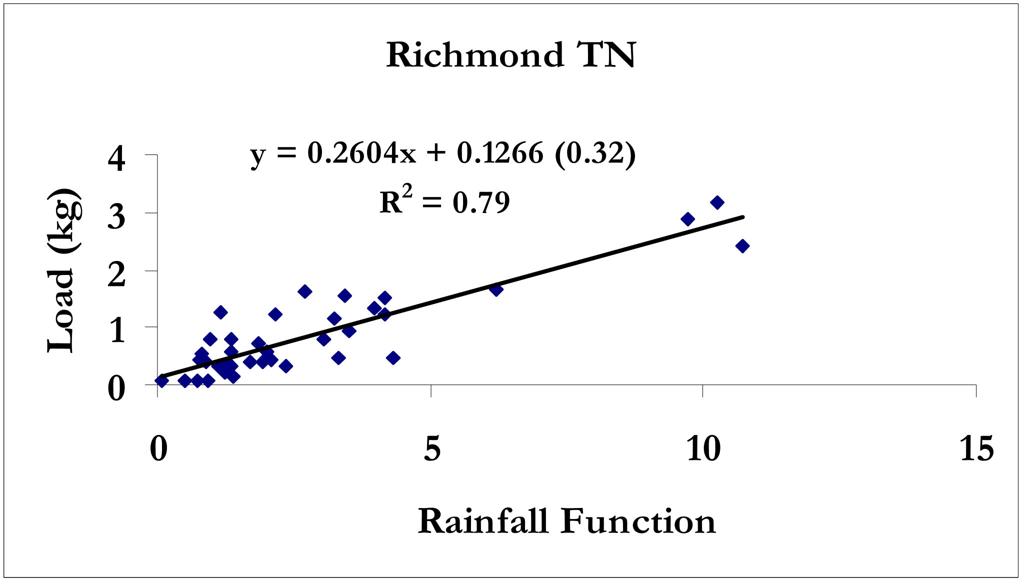

Monash Roof TP
y = 1E-05x + 1E-05 (0.00)
$R^2 = 0.28$
Load (kg)
0.0002
0.00015
0.0001
0.00005
0
0
2
4
6
8
10
Rainfall Function

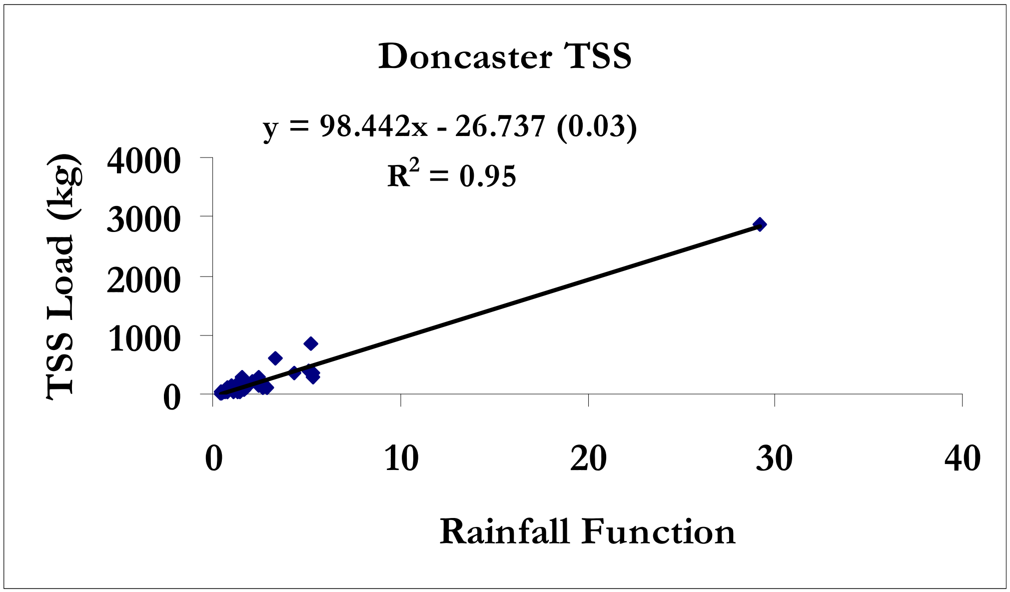

Monash Roof TN
y = 0.0001x + 0.0005 (0.18)
$R^2 = 0.37$
Load (kg)
0.003
0.002
0.001
0.000
0
5
10
15
20
Rainfall Function


Burwood East TSS
y = 22.077x + 38.092 (0.85)
$R^2 = 0.80$
Load (kg)
800
600
400
200
0
0
10
20
30
Rainfall Function

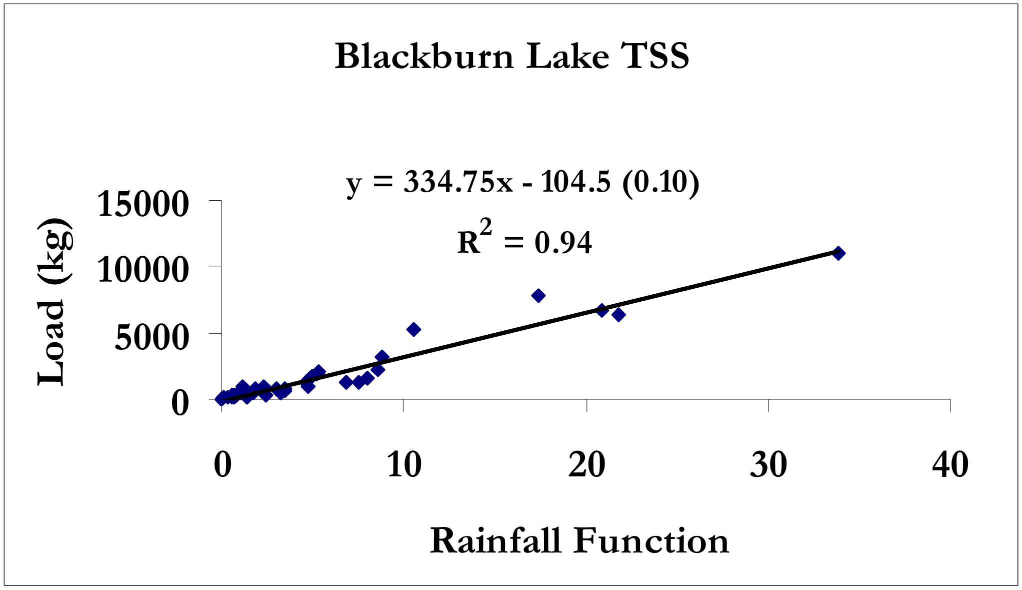

Richmond TSS
y = 97.535x + 44.411 (0.94)
$R^2 = 0.81$
Load (kg)
1500
1000
500
0
0
5
10
15
Rainfall Function

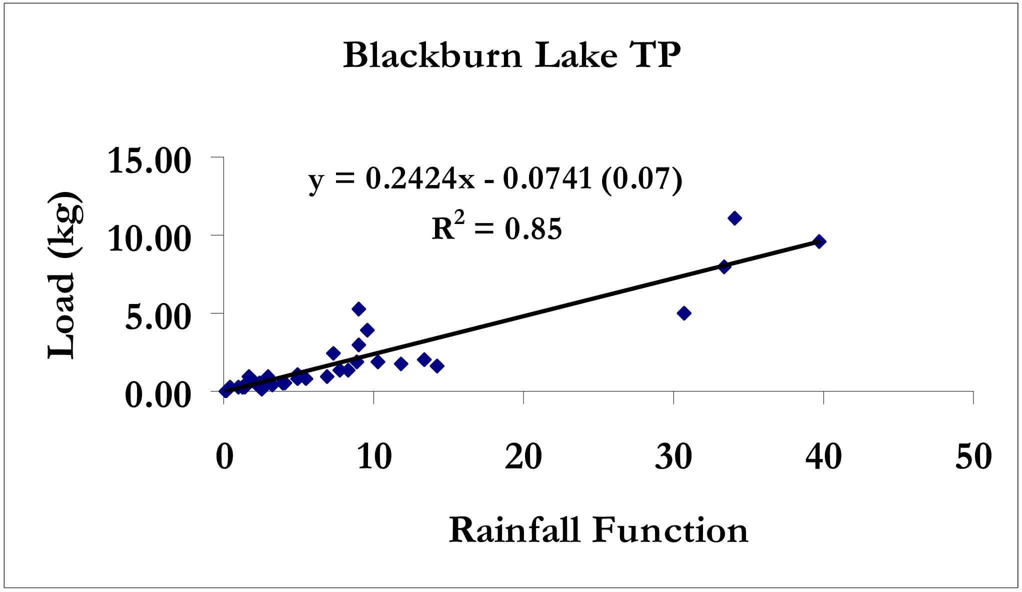

Richmond TP
y = 0.2604x + 0.1266 (0.53)
$R^2 = 0.79$
Load (kg)
4
3
2
1
0
0
5
10
15
Rainfall Function

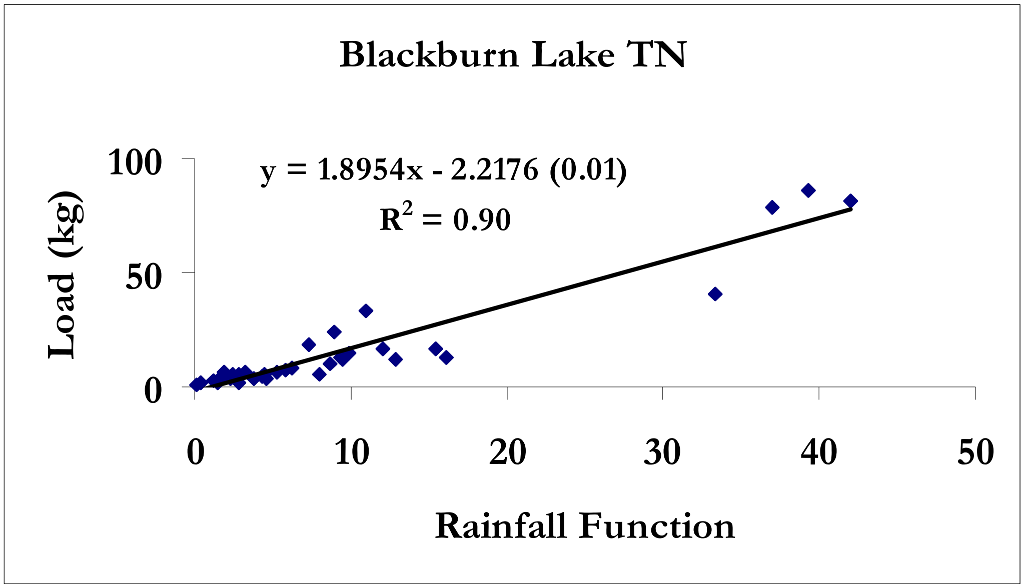

Richmond TN
y = 0.2604x + 0.1266 (0.32)
R^2 = 0.79
Load (kg)
4
3
2
1
0
0
5
10
15
Rainfall Function

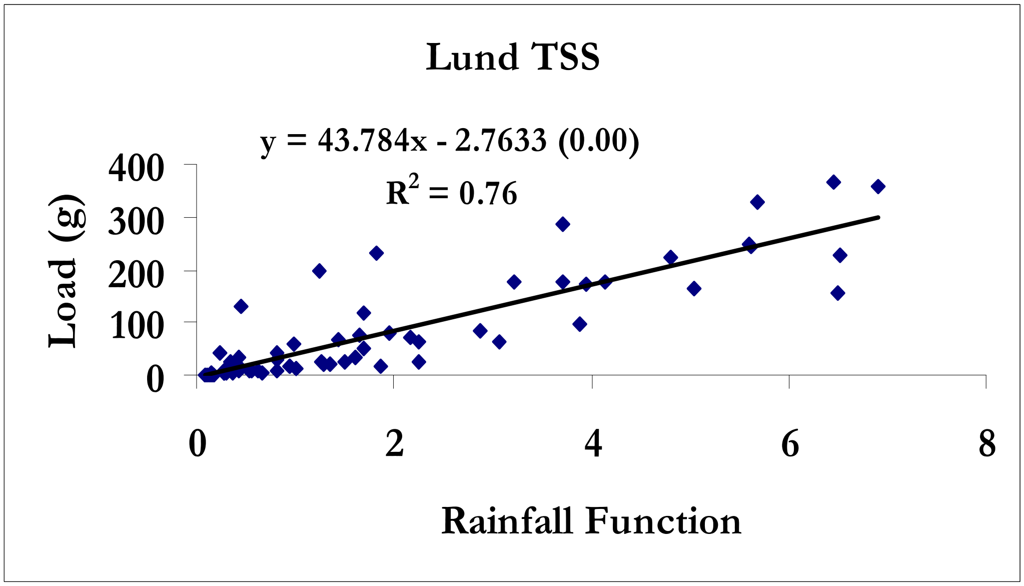

Doncaster TSS
y = 98.442x - 26.737 (0.03)
R^2 = 0.95
TSS Load (kg)
4000
3000
2000
1000
0
0
10
20
30
40
Rainfall Function

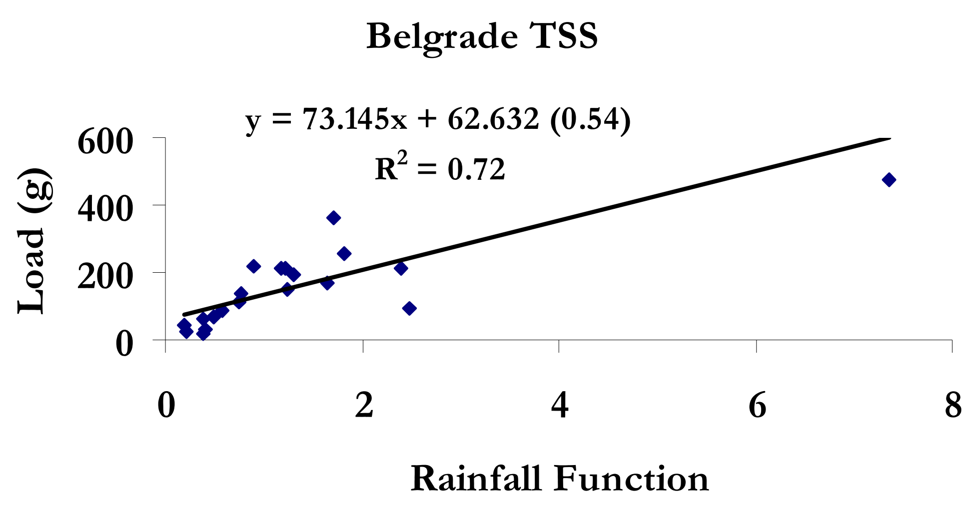
10.2 Secondary data sets (TSS, TP, TN)

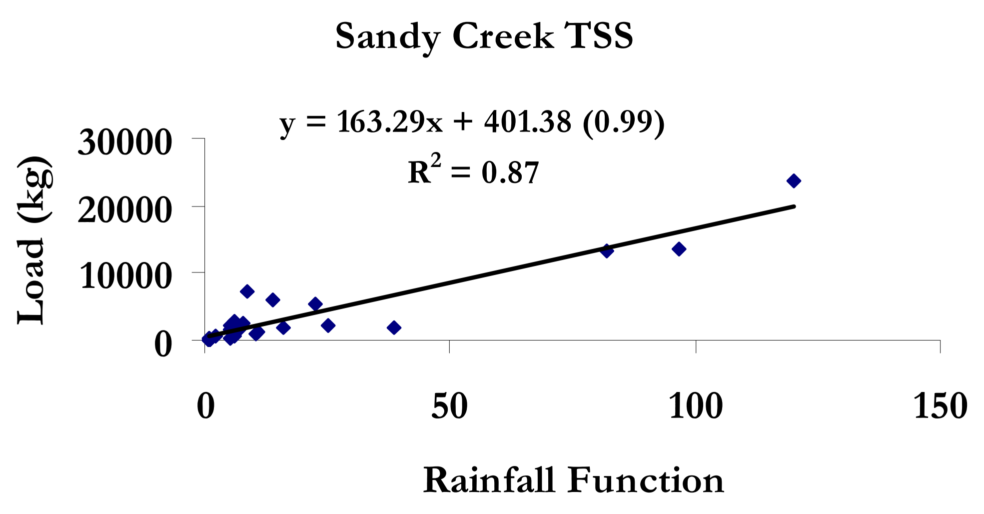


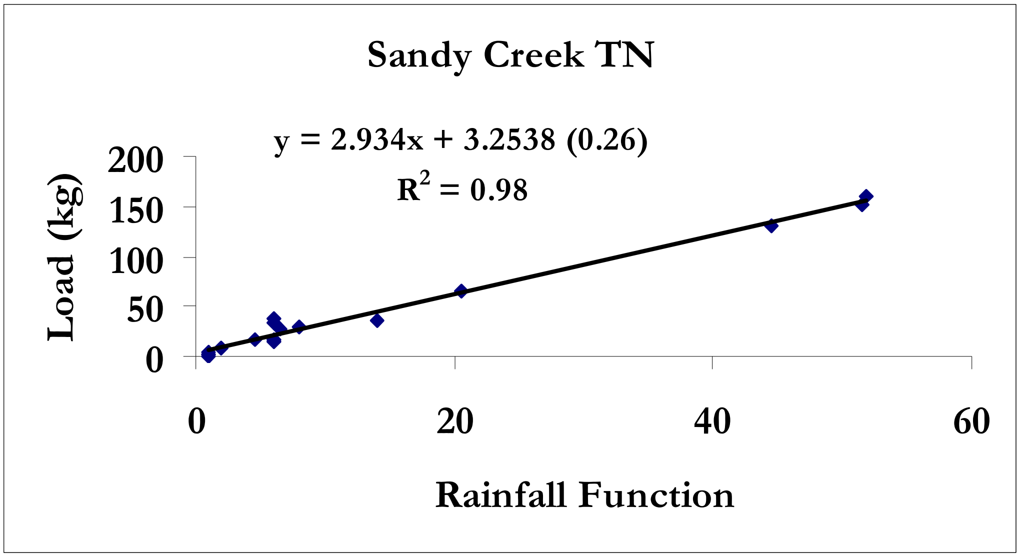



Blackburn Lake TN
y = 1.8954x - 2.2176 (0.01)
R2 = 0.90
Load (kg)
100
50
0
0
10
20
30
40
50
Rainfall Function


Lund TSS
y = 43.784x - 2.7633 (0.00)
R2 = 0.76
Load (g)
400
300
200
100
0
0
2
4
6
8
Rainfall Function

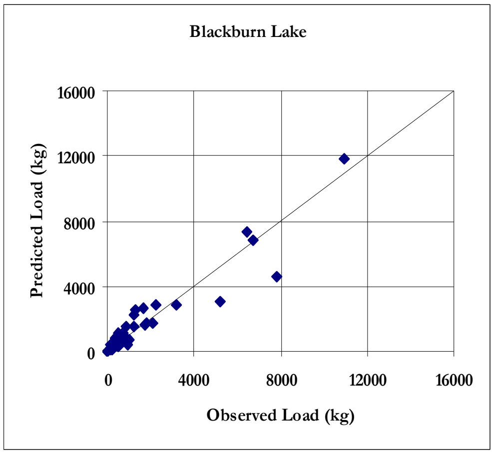

Belgrade TSS
y = 73.145x + 62.632 (0.54)
R2 = 0.72
Load (g)
600
400
200
0
0
2
4
6
8
Rainfall Function

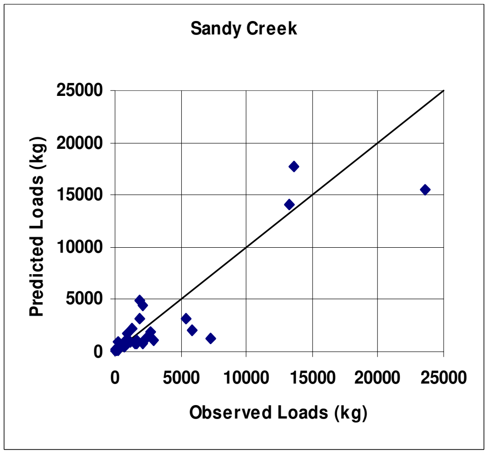

Sandy Creek TSS
y = 163.29x + 401.38 (0.99)
R2 = 0.87
Load (kg)
30000
20000
10000
0
0
50
100
150
Rainfall Function

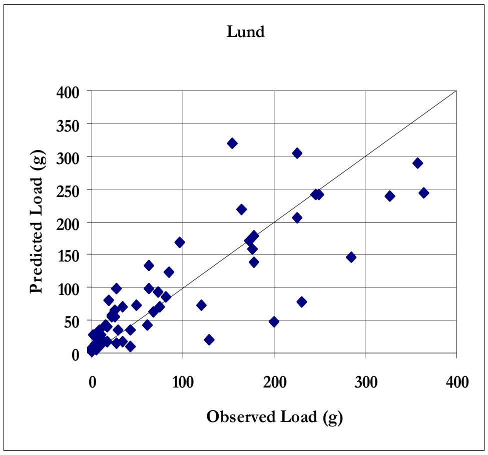

Sandy Creek TN
y = 2.934x + 3.2538 (0.26)
R2 = 0.98
Load (kg)
200
150
100
50
0
0
20
40
60
Rainfall Function

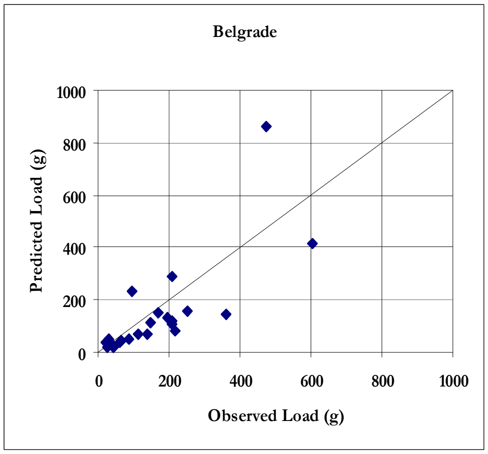
11 Appendix B – Individual catchment analysis, predicted versus observed loads

11.1 TSS – Secondary Data Sets

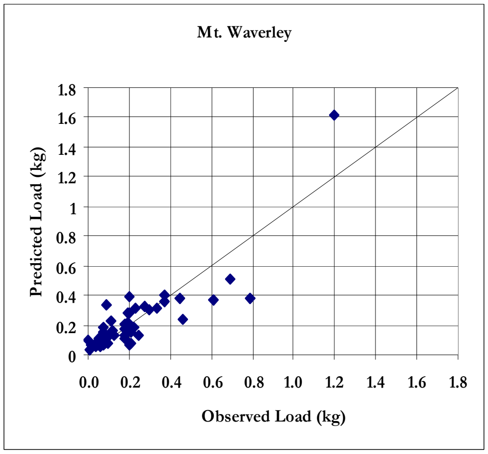


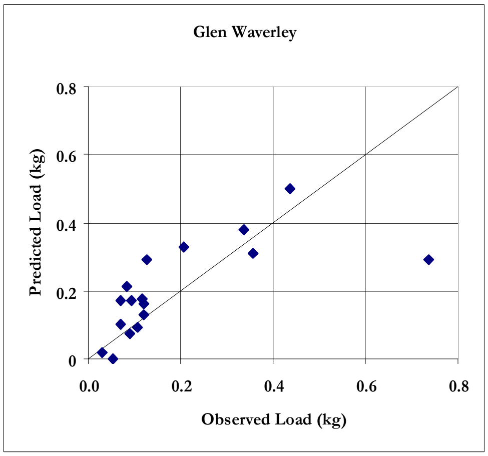


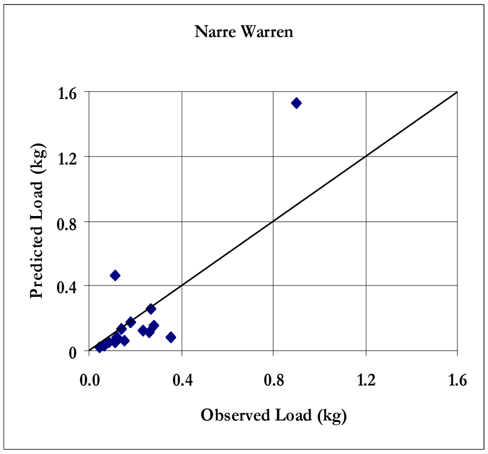


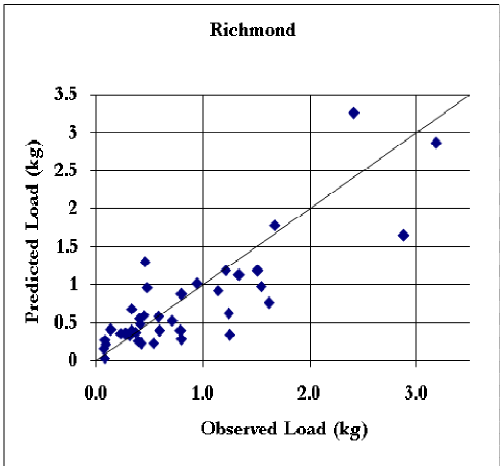


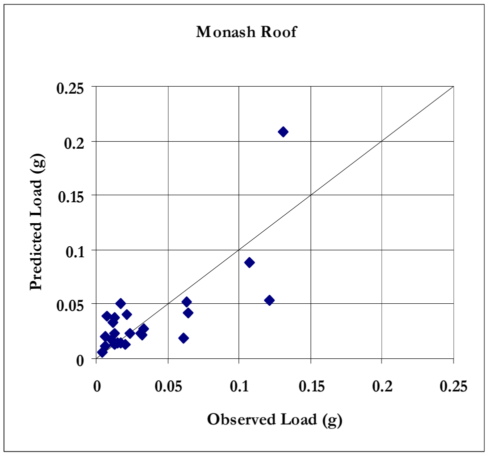
11.2 TP – Primary and secondary data sets

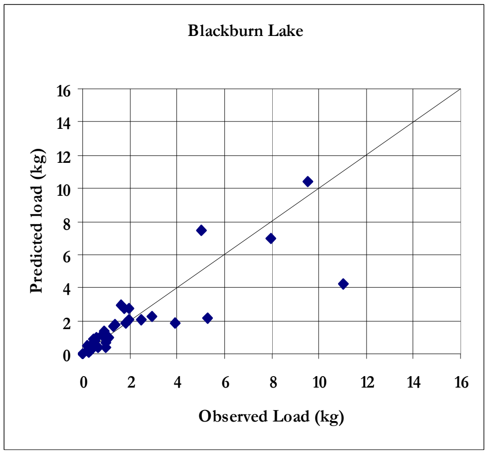


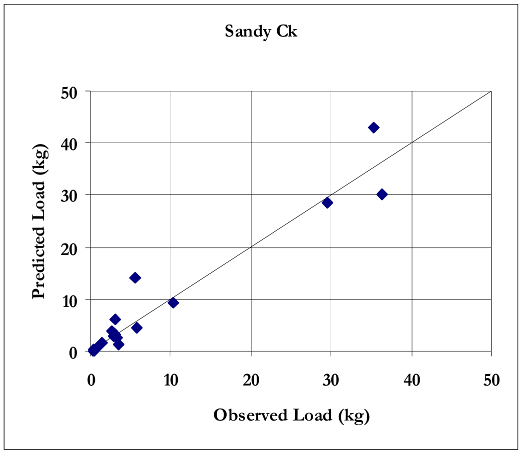




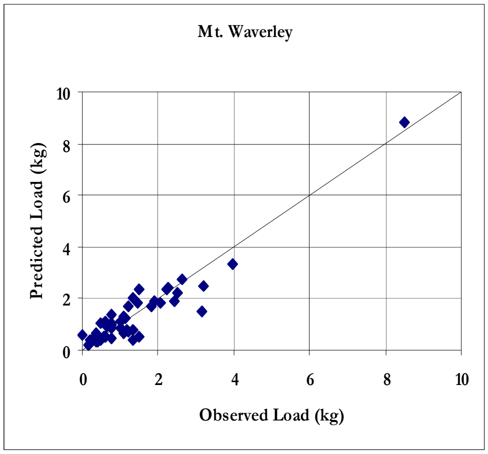


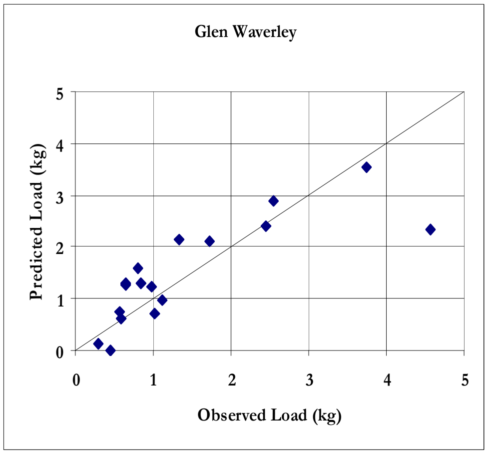


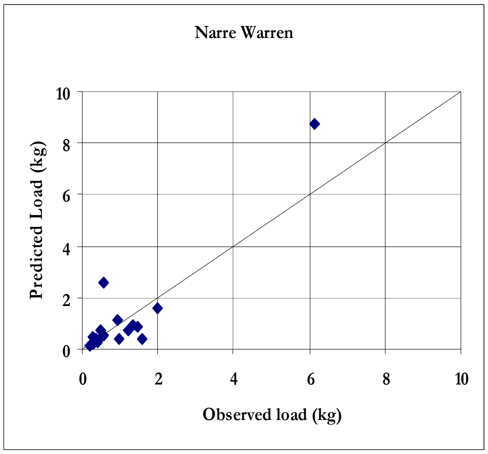


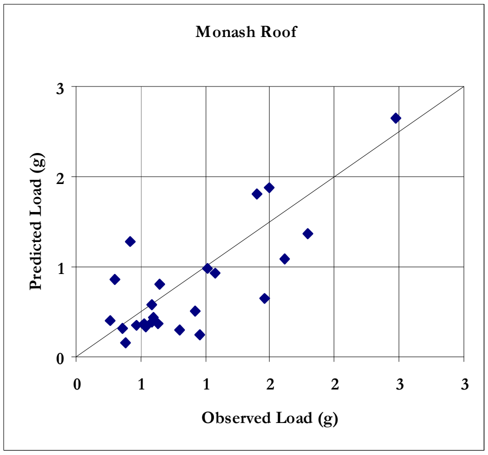

Sandy Ck
Predicted Load (kg)
Observed Load (kg)
0
10
20
30
40
50
0
10
20
30
40
50

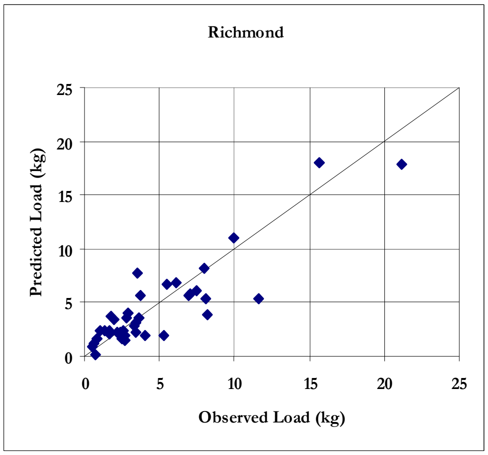
11.3 TN – Primary and secondary data sets

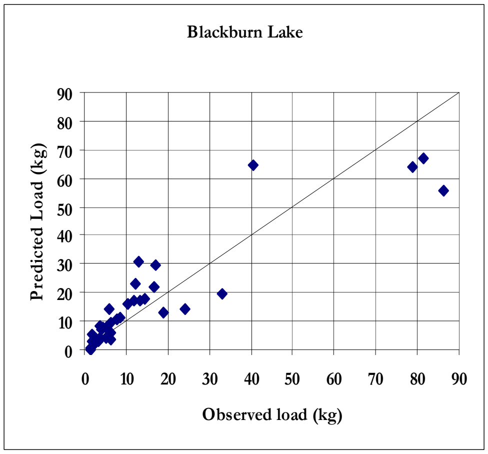


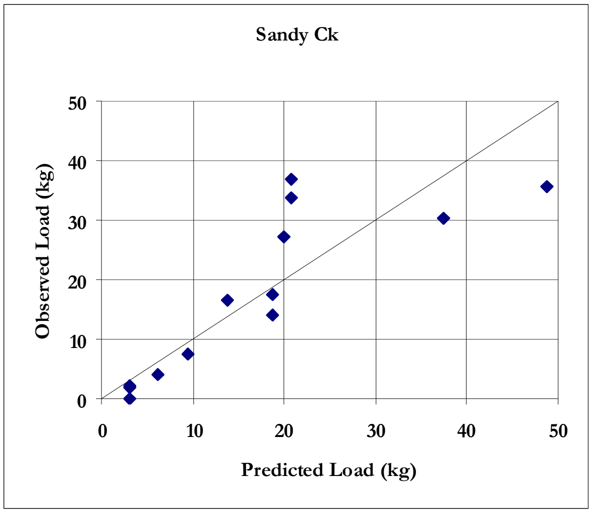




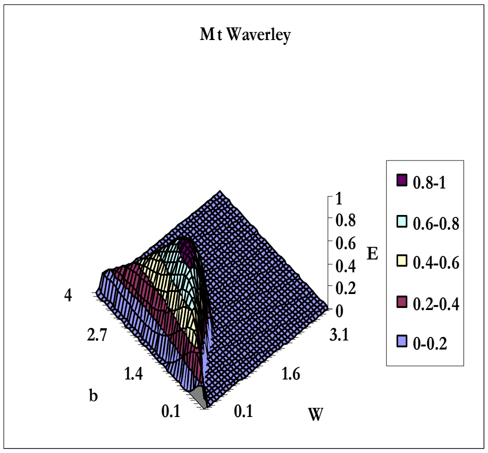


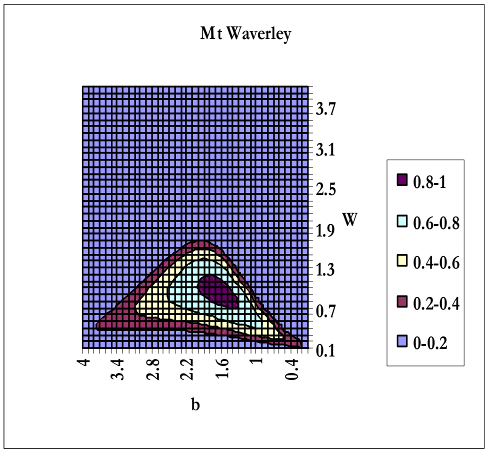


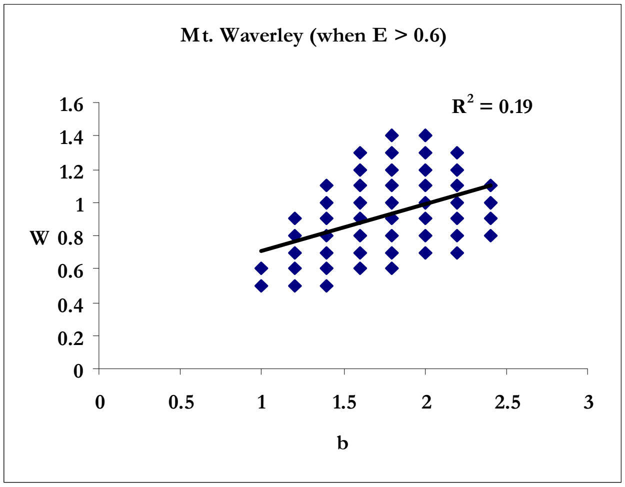


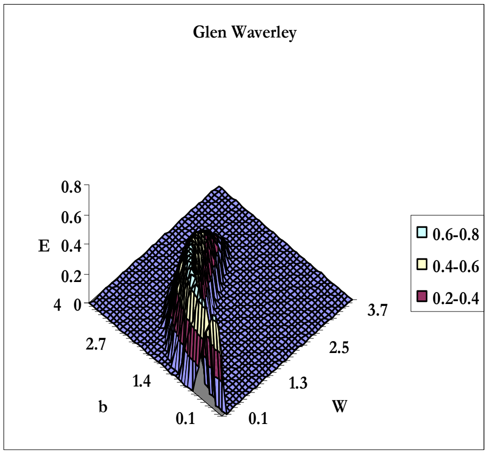

Sandy Ck
Observed Load (kg)
50
40
30
20
10
0
0
10
20
30
40
50
Predicted Load (kg)

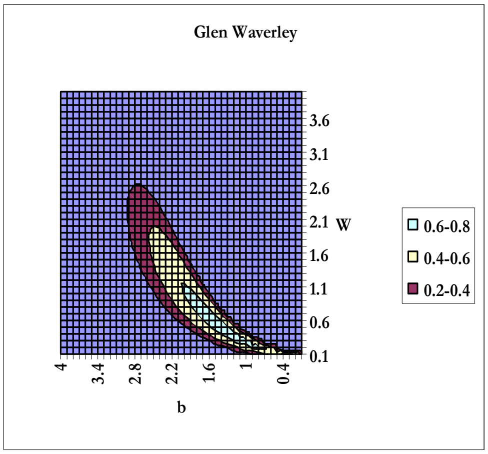
11.4 Primary data sets - TSS

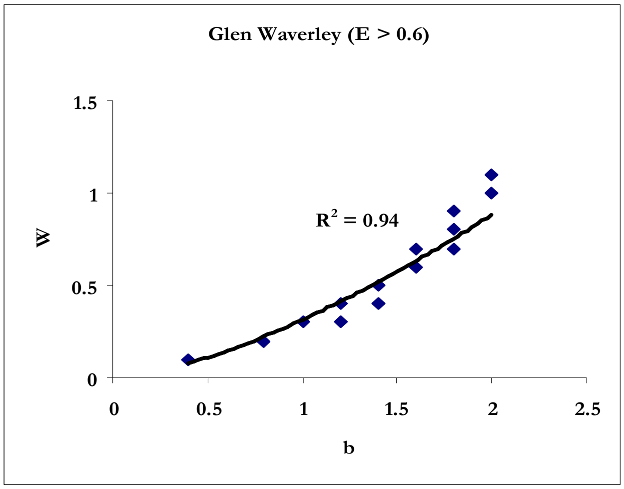


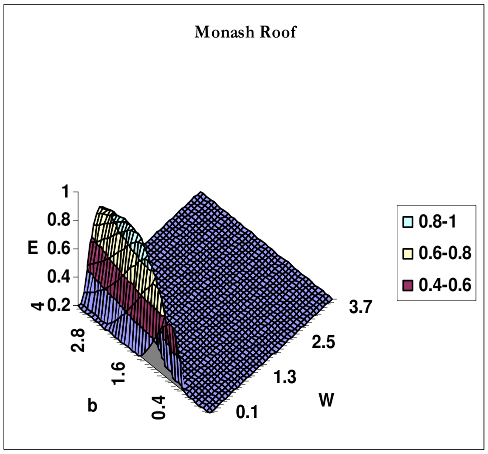


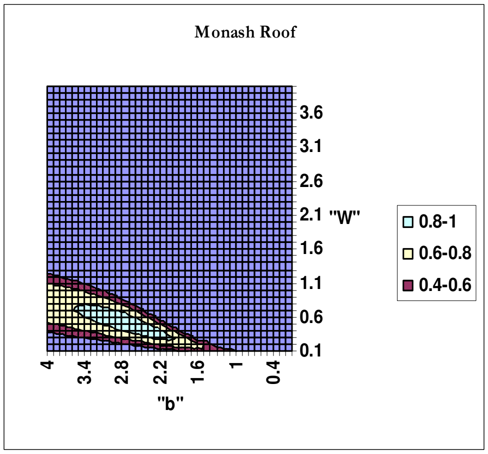


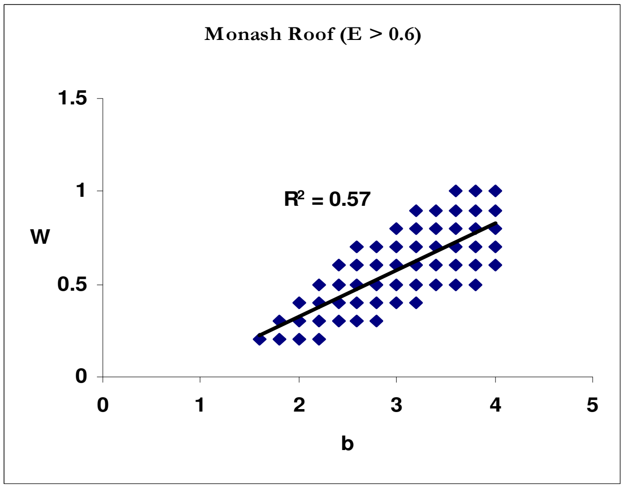


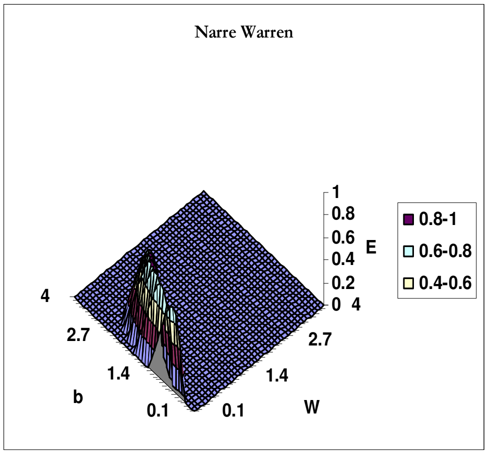


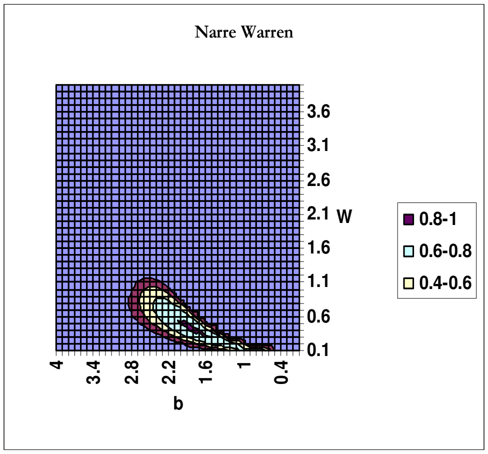

Glen Waverley (E > 0.6)

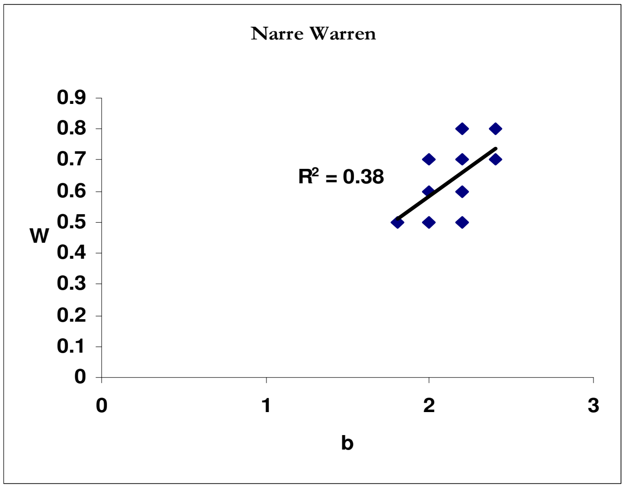

Monash Roof

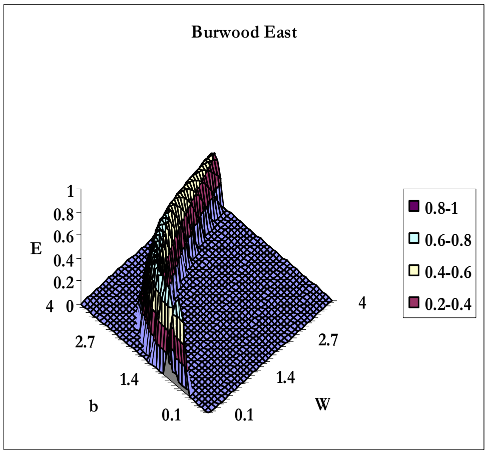

Monash Roof

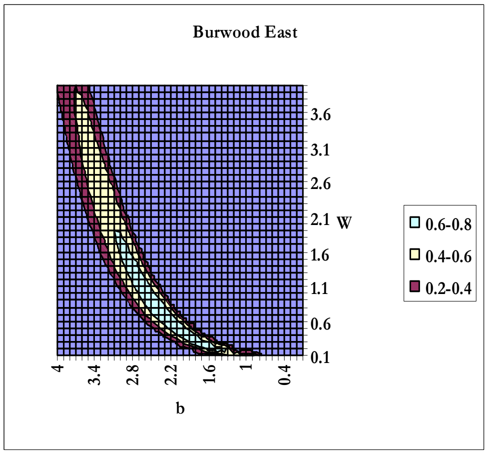

Monash Roof (E > 0.6)

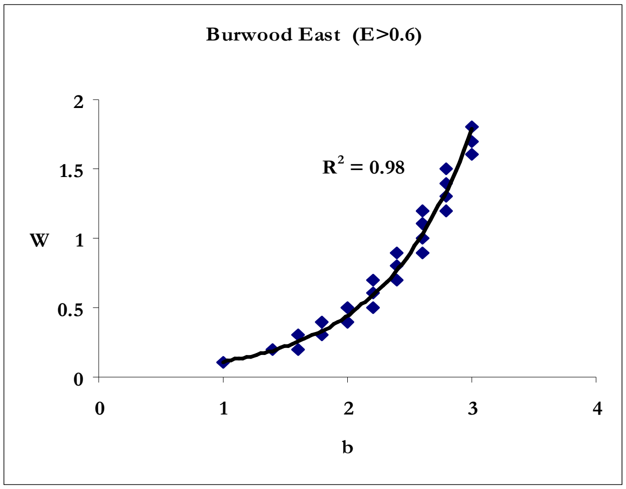

Narre Warren
1
0.8
0.6
0.4
0.2
0
E
4
2.7
1.4
0.1
b
4
2.7
1.4
0.1
W
■ 0.8-1
□ 0.6-0.8
□ 0.4-0.6


Narre Warren
3.6
3.1
2.6
2.1
1.6
1.1
0.6
0.1
W
4
3.4
2.8
2.2
1.6
1
0.4
b
■ 0.8-1
□ 0.6-0.8
□ 0.4-0.6

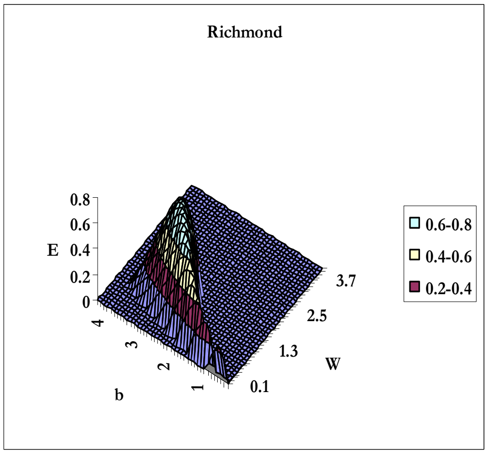

Narre Warren
W
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
$R^2 = 0.38$
0
1
2
3
b


Burwood East
1
0.8
0.6
0.4
0.2
0
E
4
2.7
1.4
0.1
b
4
2.7
1.4
0.1
W
■ 0.8-1
□ 0.6-0.8
□ 0.4-0.6
■ 0.2-0.4

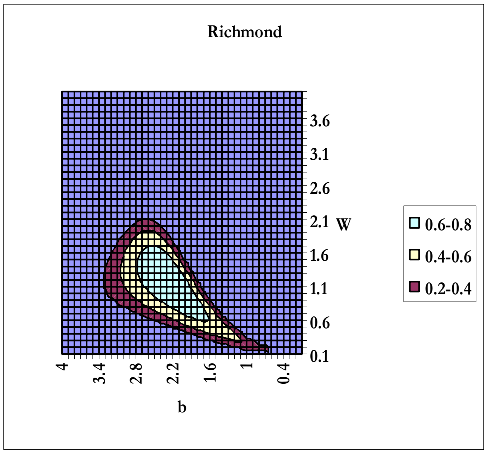

Burwood East
3.6
3.1
2.6
2.1
1.6
1.1
0.6
0.1
W
4
3.4
2.8
2.2
1.6
1
0.4
b
□ 0.6-0.8
□ 0.4-0.6
■ 0.2-0.4



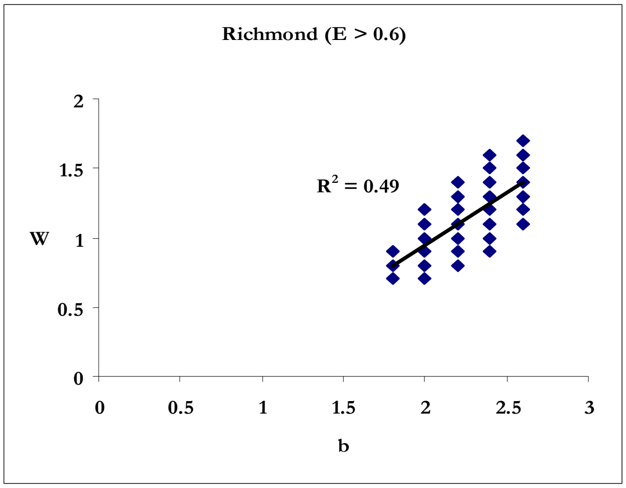
Burwood East (E>0.6)



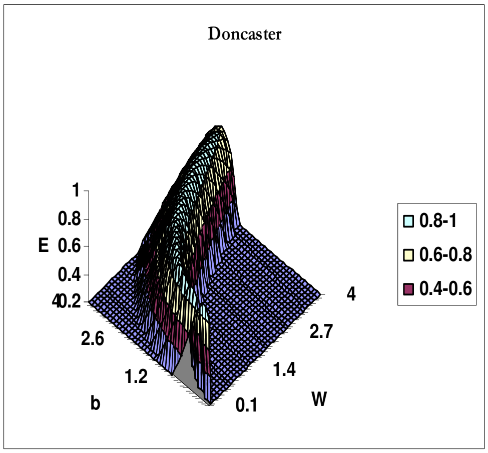
Richmond

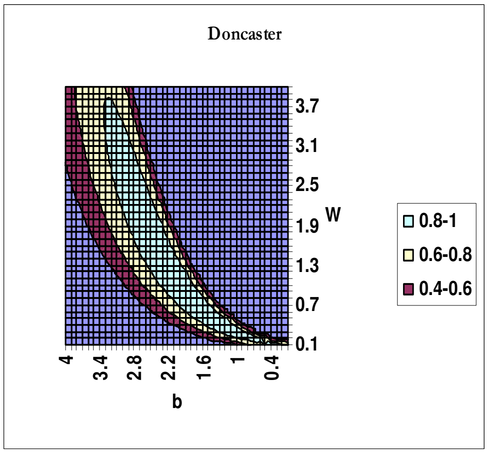


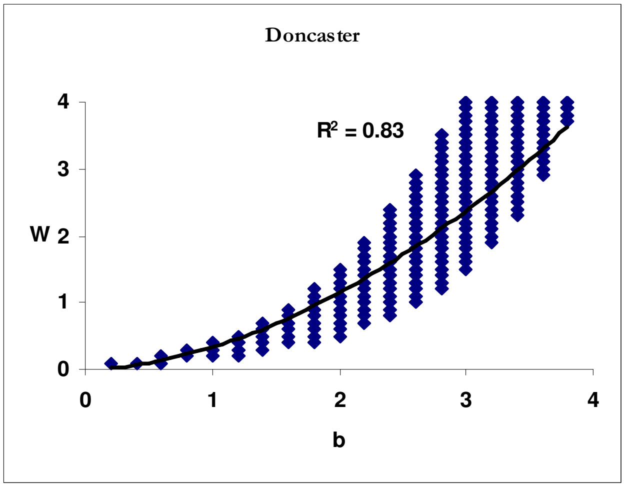
Richmond



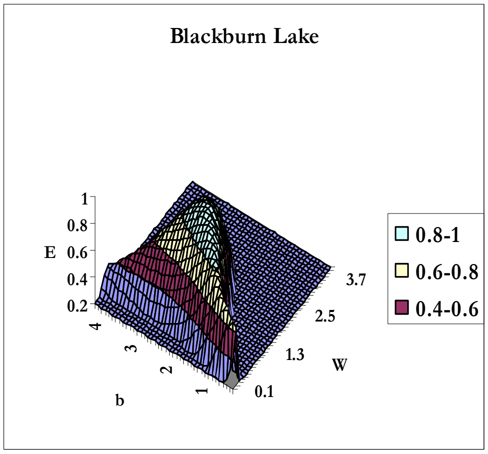
Richmond (E > 0.6)

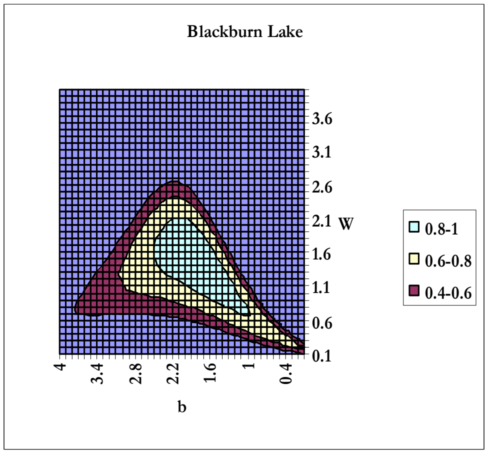

Doncaster
1
0.8
E 0.6
0.4
0.2
4
2.6
1.2
b
0.1
1.4
2.7
4
W
0.8-1
0.6-0.8
0.4-0.6

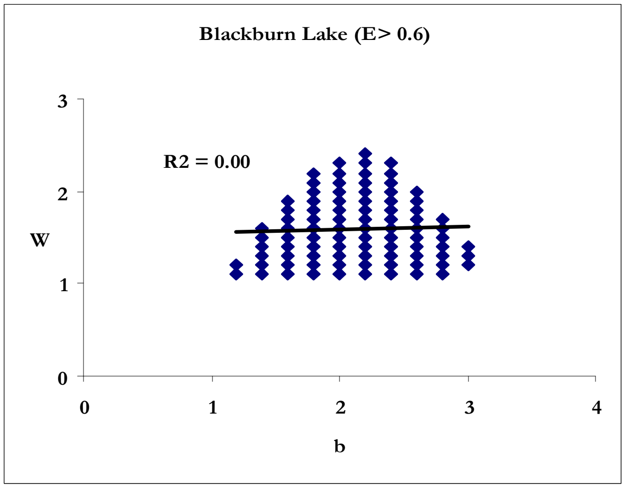

Doncaster
3.7
3.1
2.5
1.9
1.3
0.7
0.1
W
4
3.4
2.8
2.2
1.6
1
0.4
b
0.8-1
0.6-0.8
0.4-0.6

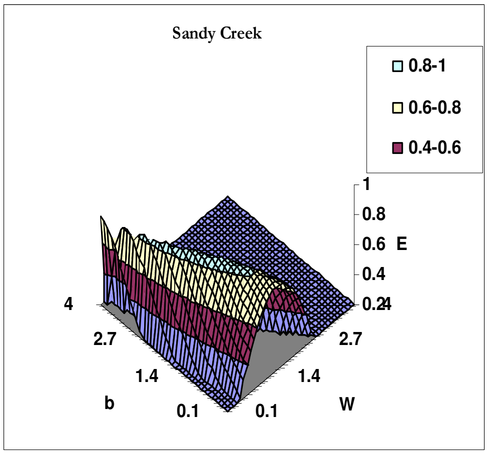

Doncaster
R² = 0.83
W
0
1
2
3
4
0
1
2
3
4
b

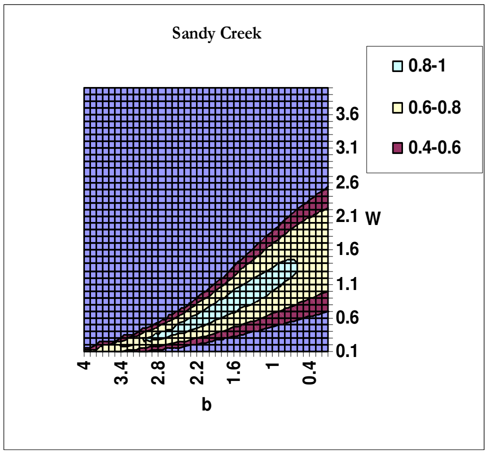
11.5 Secondary data sets – TSS

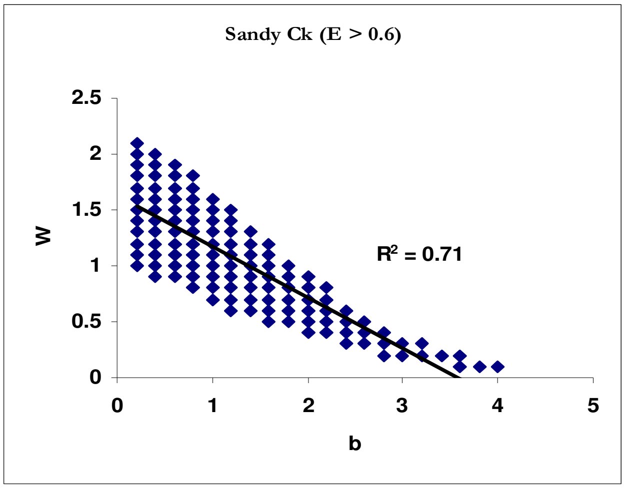

Blackburn Lake

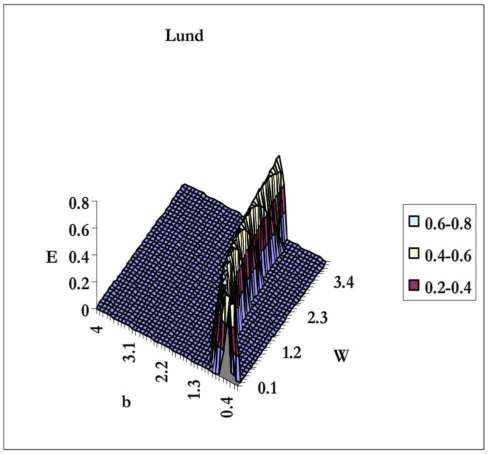

Blackburn Lake

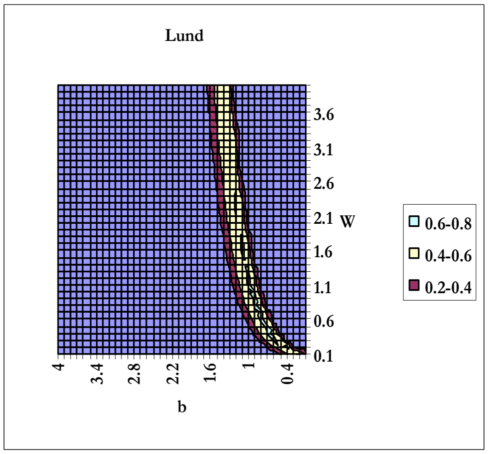

Blackburn Lake (E> 0.6)

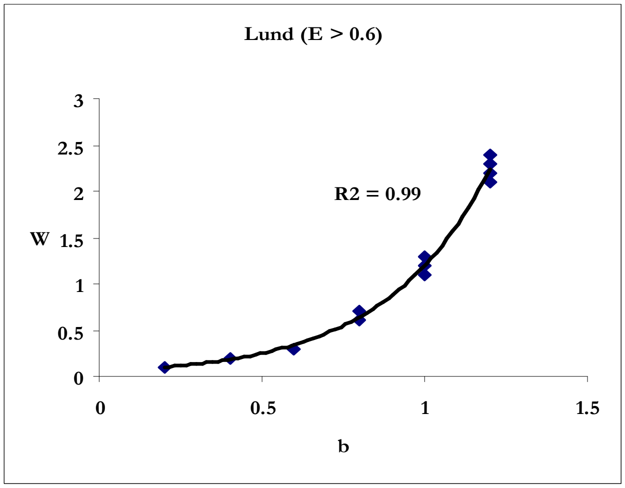

Sandy Creek

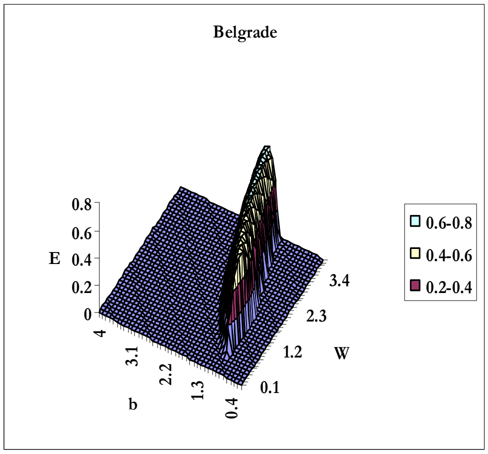

Sandy Creek

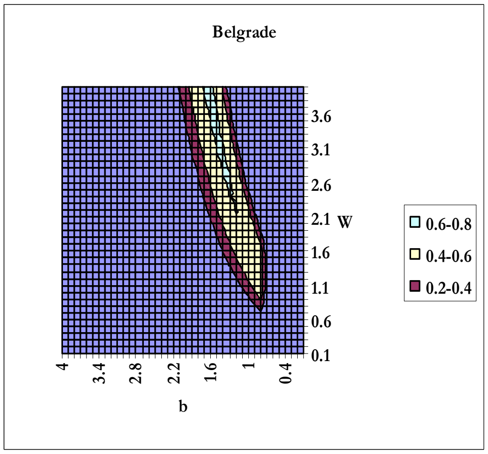

Sandy Ck (E > 0.6)
2.5
2
1.5
1
0.5
0
W
$R^2 = 0.71$
0
1
2
3
4
5
b

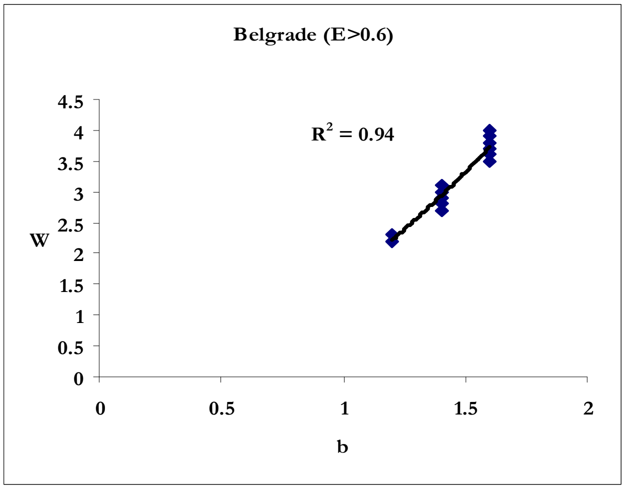

Lund
0.8
0.6
E 0.4
0.2
0
4
3.1
2.2
1.3
0.4
b
3.4
2.3
1.2
0.1
W
□ 0.6-0.8
□ 0.4-0.6
■ 0.2-0.4


Lund
3.6
3.1
2.6
2.1 W
1.6
1.1
0.6
0.1
4
3.4
2.8
2.2
1.6
1
0.4
b
□ 0.6-0.8
□ 0.4-0.6
■ 0.2-0.4

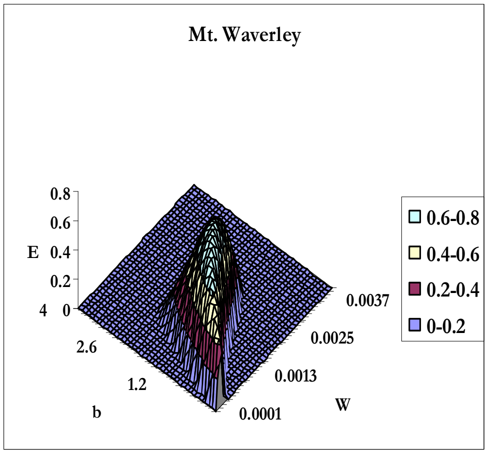

Lund (E > 0.6)
3
2.5
2
W 1.5
1
0.5
0
R2 = 0.99
0
0.5
1
1.5
b

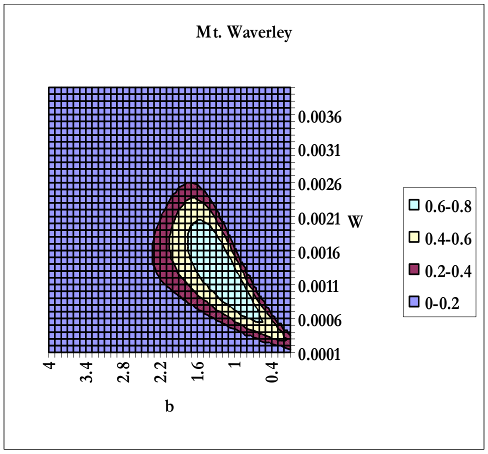

Belgrade
E
0.8
0.6
0.4
0.2
0
4
3.1
2.2
1.3
0.4
b
0.1
1.2
2.3
3.4
W
0.6-0.8
0.4-0.6
0.2-0.4

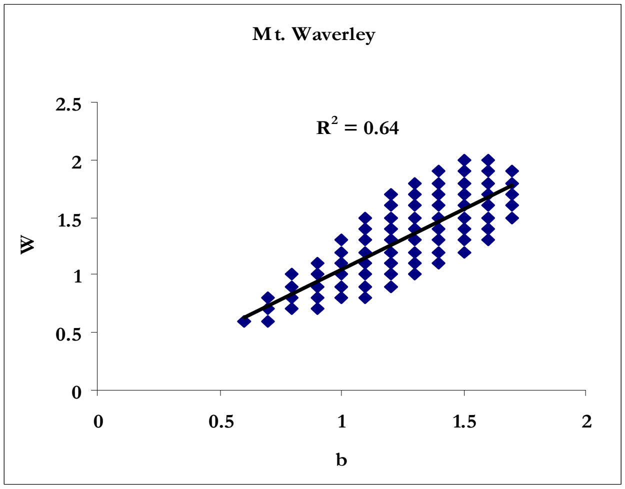

Belgrade
3.6
3.1
2.6
2.1
1.6
1.1
0.6
0.1
W
4
3.4
2.8
2.2
1.6
1
0.4
b
0.6-0.8
0.4-0.6
0.2-0.4

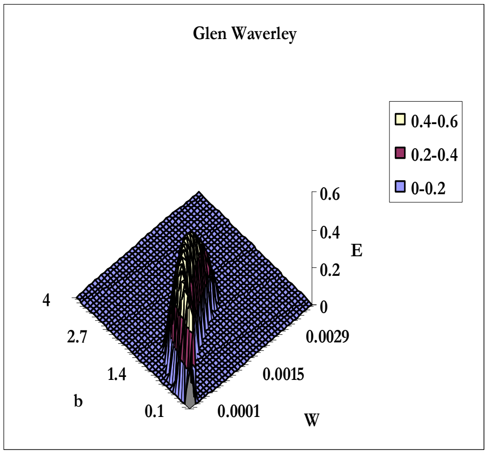

Belgrade (E>0.6)
R² = 0.94
4.5
4
3.5
3
2.5
2
1.5
1
0.5
0
W
0
0.5
1
1.5
2
b

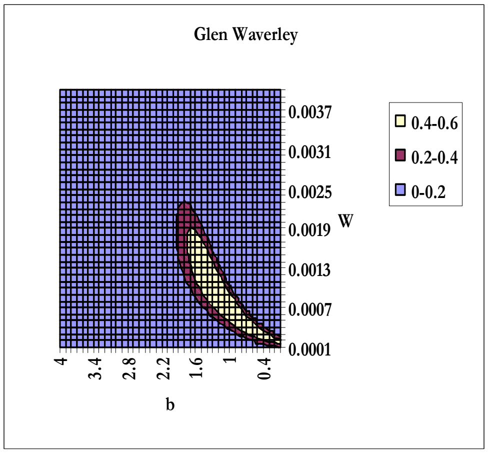
11.6 Primary data sets – TP



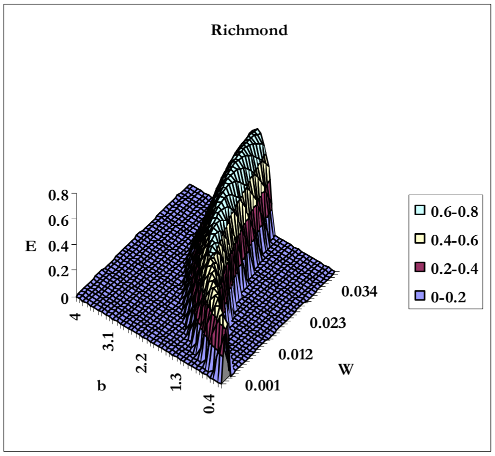


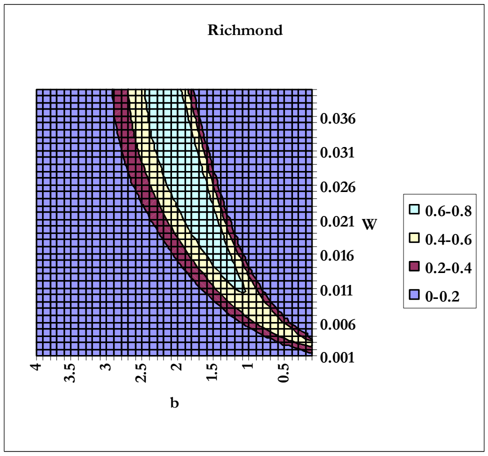


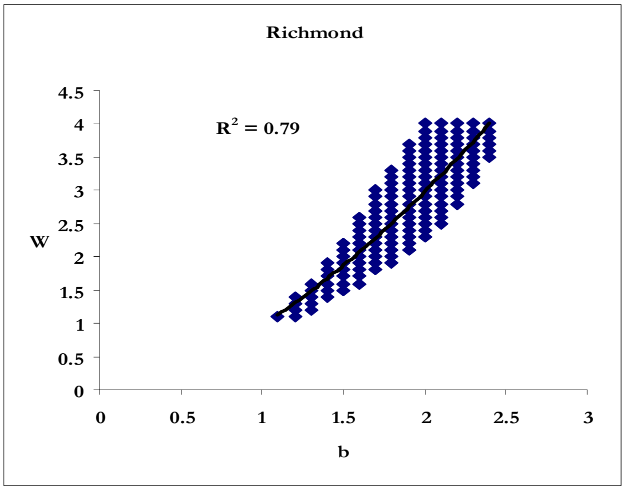




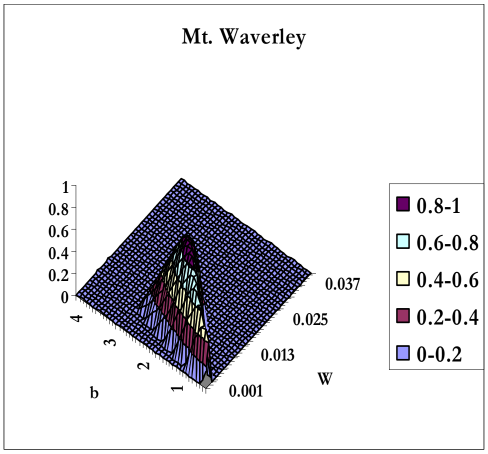
*Note, Glen Waverley had not results above E = 0.6 for TP

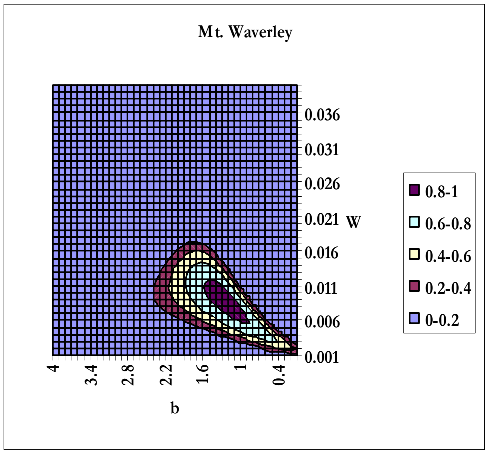


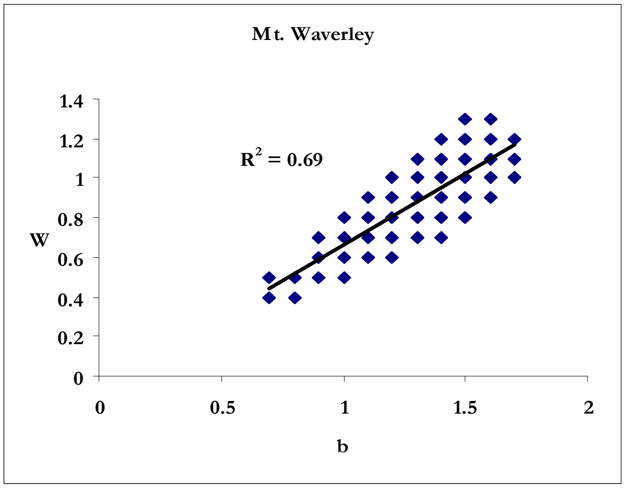


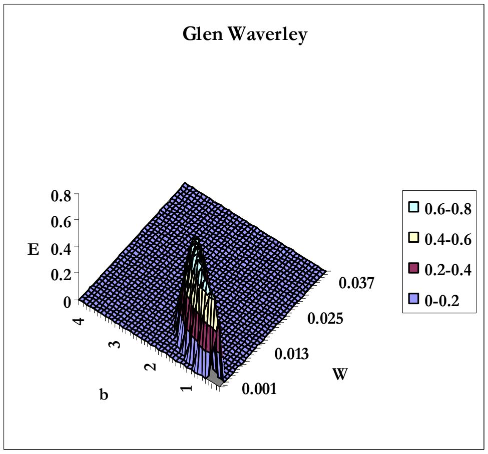


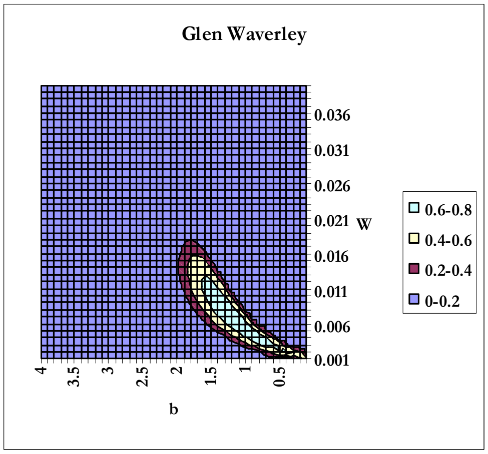
11.7 Primary data sets – TN

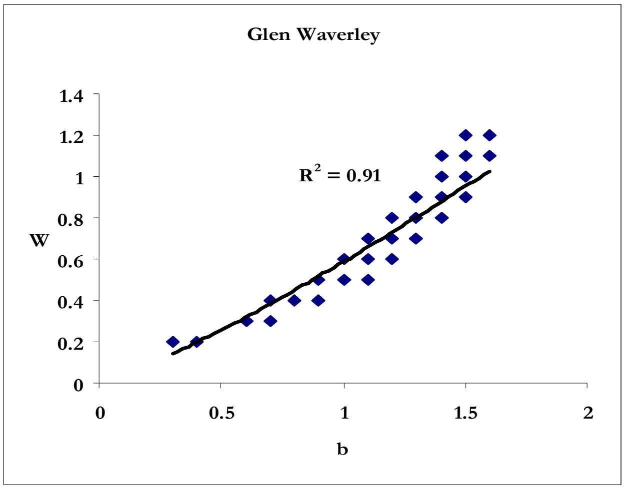


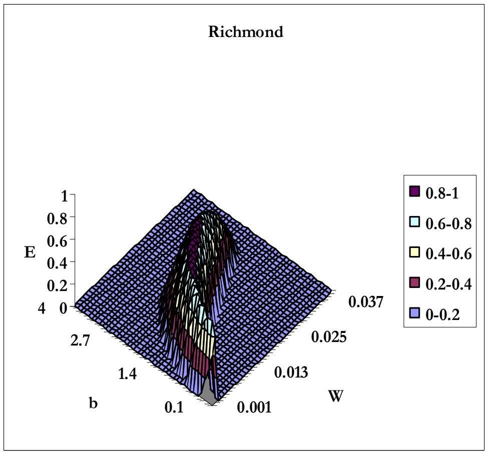


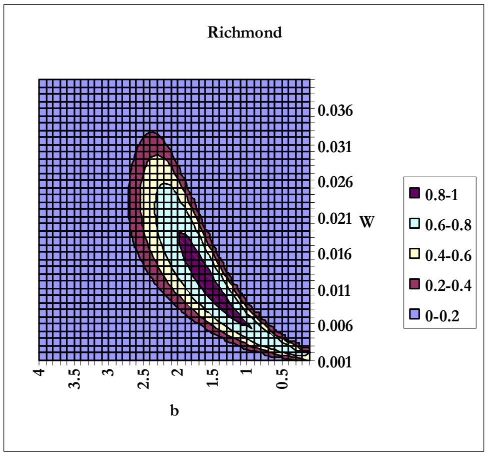


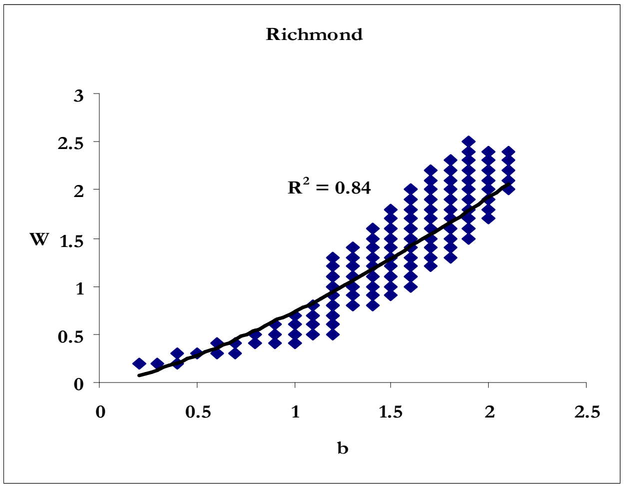





Glen Waverley
R² = 0.91
W
1.4
1.2
1
0.8
0.6
0.4
0.2
0
0
0.5
1
1.5
2
b

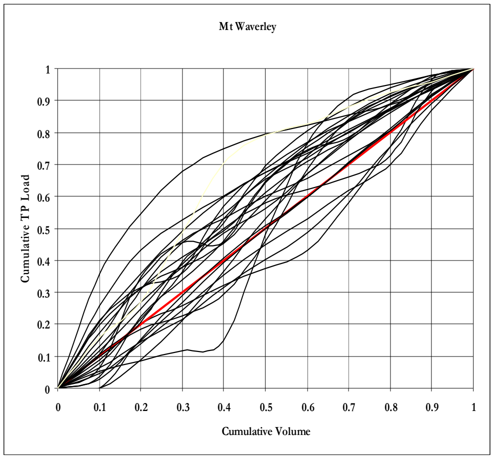

Richmond
E
1
0.8
0.6
0.4
0.2
0
4
2.7
1.4
0.1
b
0.001
0.013
0.025
0.037
W
0.8-1
0.6-0.8
0.4-0.6
0.2-0.4
0-0.2

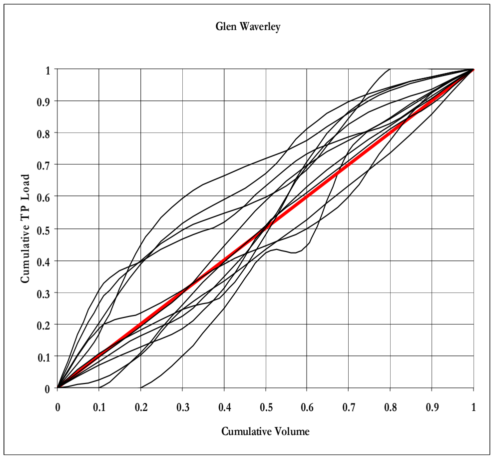

Richmond
0.036
0.031
0.026
0.021
0.016
0.011
0.006
0.001
W
4
3.5
3
2.5
2
1.5
1
0.5
b
0.8-1
0.6-0.8
0.4-0.6
0.2-0.4
0-0.2

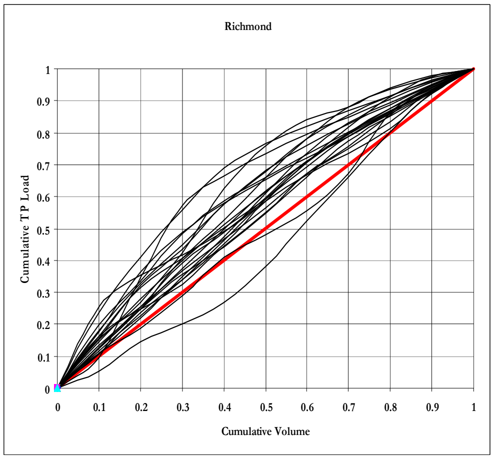

Richmond
R² = 0.84
W
3
2.5
2
1.5
1
0.5
0
0
0.5
1
1.5
2
2.5
b

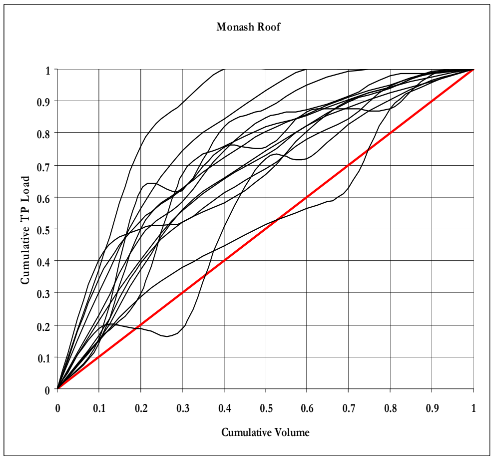
12 Appendix D – TP/TN Cumulative Load versus Cumulative Volume Graphs

12.1 TP Cumulative load versus cumulative volume graphs

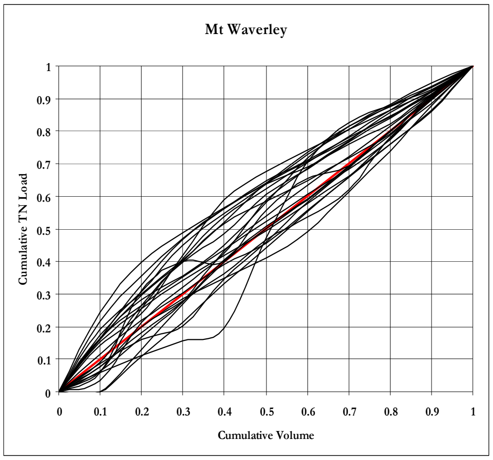


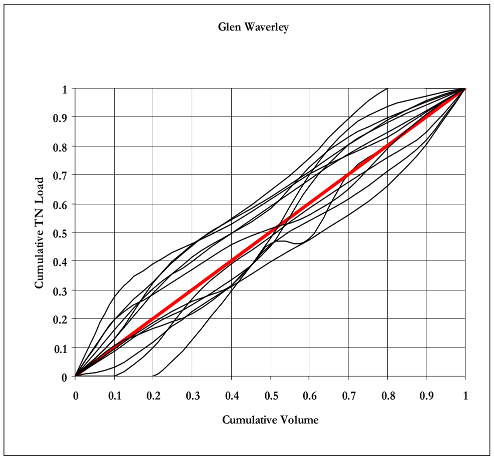


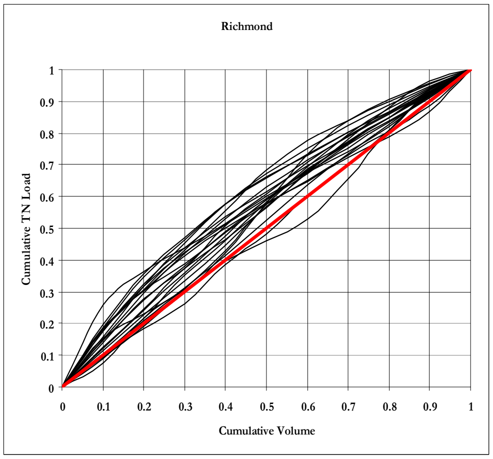


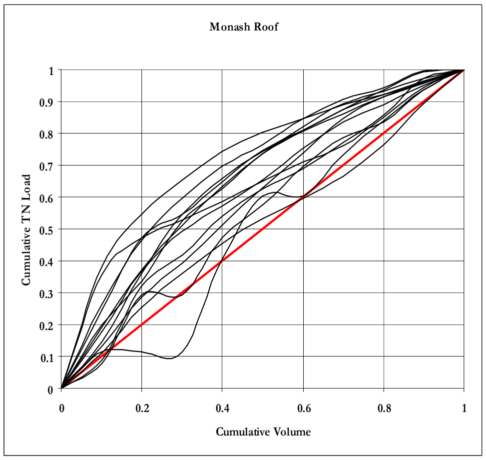


12.2 TN Cumulative load versus cumulative volume graphs









13 Appendix E - Paper published in the Journal for Environmental Engineering. April 2010 issue

New Insights into the Quality of Urban Stormwater in South Eastern Australia

AUTHOR NAMES: Matt Francey, Tim D. Fletcher, Ana Deletic, Hugh Duncan*

AUTHOR ADDRESS: Institute for Sustainable Water Resources, Department of Civil Engineering, Building 60, Monash University, VIC 3800, Australia.

AUTHOR EMAIL ADDRESS: matt.francey@eng.monash.edu, tim.fletcher@eng.monash.edu.au, ana.deletic@eng.monash.edu.au, hugh.duncan@eng.monash.edu.au

AUTHOR FOOTNOTES:

*Matt Francey Melbourne Water, 100 Wellington Parade, East Melbourne, VIC, AUSTRALIA, 3001. Tel: +61 3 9235 2221 Fax: +61 3 9235 7200

*Dr Timothy Fletcher, Director, Institute of Sustainable Water Resources, Department of Civil Engineering, Faculty of Engineering, Monash University, Tel +61 3 990 52599

*Associate Professor Ana Deletic, Institute of Sustainable Water Resources, Department of Civil Engineering, Faculty of Engineering, Monash University, +61 3 990 52940

*Hugh Duncan, Project Researcher , Institute of Sustainable Water Resources, Department of Civil Engineering, Faculty of Engineering, Monash University, +61 3 990 54981

ABSTRACT: Quantifying the quality of urban stormwater is an important prerequisite to the effective management of urban runoff, which is recognised as the major non-point source of pollution in urban areas. Although data on urban stormwater quality are widely available, they are often based on relatively limited datasets, usually containing few samples per event and/or few events per catchment. This paper reports on a large scale monitoring of the key stormwater pollutants found in urban discharges during both wet and dry weather, from seven urban catchments in South Eastern Australia. The catchments are all separately sewered (with wholly piped systems), with varying sizes and land uses. Using the same monitoring technique, between 16 and 52 pollutographs were captured at each site for TSS, TP and TN, while Event Mean Concentrations (EMCs) of heavy metals and major ions, as well as species of N and P, were recorded at a subset of sites. It was found that EMCs of TSS were around 50% less than have been typically reported in earlier literature. During wet weather, nutrients were similar to previously reported, as were most metals concentrations. However, zinc concentrations were significantly higher than previously reported. EMCs of TSS were higher during stormflows than in baseflow, whilst TN concentrations were consistently higher during baseflow. EMCs of all pollutants monitored were poorly with simple hydrological parameters (e.g. event rainfall depth), however event pollution loads correlated very well with the rainfall intensity to a power, summed over the event duration. It was not possible to distinguish an impact of land use on pollutant concentrations. The first flush effect was found not to be significant at all sites except the smallest catchment with the simplest drainage layout (the roof of a large building). All these findings have significant implication for treatment strategies, with the significantly lower than previously observed TSS requiring consideration in future modelling and treatment design.

KEYWORDS: Urban Stormwater, Water Quality, Monitoring, Total Solids, Nitrogen, Phosphorus, Heavy metals

Introduction

As direct industrial and sewage discharges to waterways have been reduced, non-point sources of pollution, such as urban runoff, have come under increasing scrutiny (Sharpin 1993; Novotny 2003). A large US EPA survey showed that urban stormwater is the most important pollutant source in US coastal waters (Burton & Pitt 2002). Similarly, stormwater has been recognised as the major source of pollution in streams and coastal waters of Australian cities (Commonwealth of Australia 2002). It is therefore not surprising that stormwater treatment systems are being built at a high rate around the world (e.g. Weiss et al, 2007).

To manage stormwater pollution we need to be able to predict pollution concentrations and loads generated from urban watersheds. Unfortunately, existing water quality models are far from being reliable, in part due to inadequate stormwater quality data sets (Bertrand-Krajewski in press; Bertrand-Krajewski et al. 1993; Deletic 1998a; Duncan 1995, 1999; Behera et al., 2006.)

Water quality data are often collected to quantify Event Mean Concentrations (EMCs); i.e. the flow-weighted mean pollutant concentrations over the duration of a storm event. EMCs are typically log distributed in urban catchments (Duncan 1999, Fuchs et al. 2004). Usually more than one storm event is monitored at a site and the site average EMC (µEMC) is then calculated (the *mean* of all measured EMCs). . The Site Mean Concentration (SMC) may also be calculated, by dividing the total mass of pollutants across all monitored events by their total runoff volume. The SMC is typically used for estimating annual pollutant loads.

Major French and US studies in the early 1980s enhanced understanding of stormwater quality and its pollution impacts (Athayde et al 1983; Deutsch & Desbordes 1981). Further monitoring programs, particularly in the USA (Pitt & Maestre 2005; Smullen et al. 1999), have progressively presented more comprehensive and precise data. However, the principal factors behind the generation of stormwater pollutant loads remain contentious (Francey et al. 2005; Vaze & Chiew 2003).

Studies in Australia have usually been undertaken by management agencies (Gutteridge Haskins and Davey & Environment Protection Authority of Victoria 1981; City Design 2003) and have not specifically focussed on reducing the uncertainty involved in measurement. Australian monitoring programs with rigorous quality control have been very few, and are usually small in size, focusing on 1-3 catchments or small plots (Cordery 1997, Ball 2000; Vaze & Chiew 2002, 2003b).

However, several large compilations of world wide data sets have been conducted (Duncan 1999; Fuchs et al. 2004; Smullen et al. 1999), attempting to differentiate stormwater quality based on land use, region, and season. In some cases there was a tendency towards higher pollutant concentrations for industrial and commercial sites than residential. Another important finding was that concentrations of many (but not all) pollutants are higher during storm discharges than during dry-weather flows. Such studies have also found that solids are the major carriers of pollution with on average 70% of heavy metals and phosphorus being attached to fine particles (Deletic & Orr, 2003; Duncan 1999). Table 1 shows typical concentrations of pollutants in stormwater from a recent compilation study (Fuchs et al. 2004).

In another compilation, Smullen et al. (1999) analysed stormwater quality data collected from several USA sampling programs conducted over the past 20 years, comparing the results with the original US NURP results (Athayde et al 1983). Importantly, they showed that stormwater quality is *significantly better than previously reported* (Table 2). For example, EMCs of TSS were on average half of the reported values 20 years ago. Similar findings have been made by other researchers (Pitt & Maestre 2005; Smullen et al. 1999). However, the reasons for this are poorly understood (apart from the drop in Pb being explained by the removal of leaded petrol), with some tentative hypotheses put forward. For example, it was suggested that past monitoring practice was biased toward bedload sediment collection, with advances in sampling systems resulting in more ‘representative’ samples (Smullen et al. 1999).

It seems timely to re-evaluate stormwater pollution, since most stormwater management objectives and design of treatment systems are based on previously published stormwater pollution concentrations (without any local measurements). Confirmation of suggested differences in recent stormwater pollutant concentrations could result in significant differences to practice and design standards. This and the relative lack of good quality data for Australian conditions have led to the development of a large scale monitoring program, with three primary aims:

1. To collect high definition data on urban stormwater quality for South Eastern Australia, in order to provide reliable estimates of current pollutant concentrations, and compare with previous estimates of pollutant concentrations;
2. To test the principal factors behind pollution generation.
3. To provide practical guidance on how this information may improve current stormwater treatment practice.

14 Methods

Catchments and Monitoring Program

The seven catchments in this study are located in the south-eastern suburbs of Melbourne, Australia, and were chosen to reflect a variety of land uses and catchment sizes (Table 3). All are serviced by separate underground stormwater systems and none of the catchments contains significant stormwater quality treatment. Melbourne has a temperate climate, with warm summers and cool winters. Rainfall occurs throughout the year, with a slight spring maximum. Mean annual rainfall across the area of interest ranges from 600 to 800 mm/year. The sampling was undertaken between November 2003 and December 2005, rainfall was slightly below the long term average over this period.

Rainfall data was collected using 0.2 mm tipping gauges compiled at a one-minute timestep, located as close as possible to the catchment centroids.

Doppler-based flow-meters (Sigma 950) were installed in pipes close to the catchment outlets. The runoff rate was recorded at one-minute intervals using the assessed cross-sectional area and the measured flow velocity. Flow-meters were positioned 2 – 4 centimetres above the pipe invert to minimise backwater effects and clogging of the sensor. Inspection and cleaning of the sensors was undertaken after each storm event, and on a fortnightly basis during periods without storms.

At each site, Sigma 900 auto-samplers were installed alongside the flow-meters, and programmed to start sampling after the flow rate reached a trigger value (which varied depending on the catchment). Subsequent samples were taken after predetermined volumes passed through the pipe (i.e. a flow-weighed sampling approach was used). The method was refined during a preparatory period by adjusting the volumes to the runoff rates experienced at each site. In very large events the bottles were manually replaced when exhausted – therefore allowing more than 24 samples to be taken for these storms. Prior to sending samples for chemical analysis, a number of rules were followed to ensure consistency:

- Rainfall was: (a) >0.6mm and (b) recorded from the *time of concentration* prior to the first sample being taken, to the *time of concentration* before the last sample (an average time of concentration was estimated for each catchment based on the time between recorded rainfall and the response recorded in the outflow).
- Runoff was: (a) recorded between the first and last samples and (b) have a volumetric runoff coefficient <1.
- Time between events was at least 4 hours (approximately 8 times the time of concentration of the largest catchment).
- The number of samples was at least five, with a sample interval of no more than 2 hours.

Dry weather pollutant samples were taken after a period of at least three days without rain, by simply holding a sample bottle in the flow at the bottom of the pipe or at the pipe outlet of each catchment. The Monash Roof was the only site without a continuous baseflow.

All water samples were sent to a Australian National Association of Testing Authorities (NATA) accredited laboratory to be analysed for Total Suspended Solids (TSS) via measurement of sediment mass, and Total Phosphorus (TP) and Total Nitrogen (TN) via colorimetric Flow Injection Analysis (FIA), all according to Standard Methods (Greenberg et al. 1999). For most events, all samples were analysed, whilst for some events, flow-weighed composites were made and only Event Mean Concentrations were determined. For the same composites, speciation analysis was undertaken to determine Ammonia Nitrogen (NH_3), Filterable Reactive Phosphorus (FRP), Oxidised Nitrogen (NO_x) and Total Dissolved Nitrogen (TDN), again using FIA according to Standard Methods. Samples for total heavy metals and major ions (stored in acid-washed bottles) were acid-digested using concentrated nitric acid, and concentration determined using ICP Emission. The device detection limits were 0.7 µg/L for Cd, 0.3µg/L for Cu, 0.6µg/L for Pb and 0.5 µg/L for Zn.

15 Number of Monitored Events

To achieve reliable estimates of pollutant concentration at the least cost, an uncertainty analysis was undertaken to determine the number of events which must be sampled, to reliably determine SMCs or site average EMCs for each catchment (Francey *et al.*, 2004). We calculated uncertainty as the 95% confidence interval of the mean. In order to set a "target" uncertainty, we use the uncertainty which occurs typically occurs for a single sample, due to a combination of sampling and laboratory analytical errors. This 'single sample uncertainty' has been previously found to be around 30% for TSS (Ahyerre *et al.*, 1998; Bertrand-Krajewski *et al.*, 2002). Therefore, if enough events could be sampled to reduce the uncertainty (95% confidence interval) of the SMC and site average EMCs to approximately 30% of the mean value, we could be confident that the monitoring program was providing enough information to describe the SMC and site average EMC (µEMC) with a realistic minimum uncertainty. Although the 30% measurement uncertainty has been reported mainly for TSS, the same value has been used as a guide for TP and TN, as most of the uncertainty is in the sampling methodology, rather than in the laboratory analysis (Bertrand-Krajewski & Bardin, 2002b), and is therefore not highly pollutant dependent.

An independent data set was used for this pilot-analysis, collected in 1996-1997 in a 221 ha residential catchment (Blackburn, total fraction impervious 0.58), also located in eastern Melbourne. The EMCs of TSS, TN, and TP were available for 34 events at this catchment. Bootstrapping (a method of subsampling with replacement, see for example Chernick, 1999) was used to determine how the number of events sampled (N) impacts on the uncertainty of SMC and µEMC for each pollutant. The advantage of the bootstrapping method, as opposed to Monte Carlo methods based on assumed distributions of the data, is that it can be applied to any distribution, not just a known distribution such as normal or log-normal, for example. Random subsamples of n = 5, 10, 20, 30, 40, 50…100 events were taken from the Blackburn data set (replicated 1000 times) and the SMC and μEMC were calculated for each case. Using bootstrapping (sampling with replacement) therefore allowed us to calculate the uncertainty for values of n higher than the number of events (34) measured at Blackburn Lake. In doing so, we are relying on the assumption (Chernick, 1999) that the sample parameters of n samples represent the characteristics of the population.

The 95% confidence interval was then determined using all 1000 values for each tested n, and plotted against n. The required n (ie. the number of events to be sampled in the monitoring campaign) was thus determined as being when the uncertainty (95% confidence interval) became smaller than 30% of the SMC and μEMC. The analysis showed that 50 events are needed for TSS, and 25 for TP and TN.

As we had to collect 25 more events to characterise TSS than the other pollutants, it was decided that composite-sampled EMCs of TN and TP would also be measured in these events. However, technical problems and unusually dry conditions resulted in the number of events sampled falling short of the target in some instances (Table 4).

As we were not able to conduct similar analyses on other pollutants such as heavy metals (due to a lack of an available pilot data set), an attempt was made to collect as many EMCs of heavy metals as resources and time allowed. Between 8 and 17 EMCs were collected for three catchments as shown in Table 5.

Data Analyses

Data Preparation - The total event rainfall, the average rainfall intensity over the event, and the maximum rainfall intensity in a six-minute timestep were calculated for each event. It is important to note that the rainfall intensity within a timestep was calculated by assuming that total rainfall depth recorded in one time step occurred just within that very time step. Therefore, the recorded intensity depends strongly on the temporal resolution of rainfall records. A six-minute timestep was used as it is widely used in Australia and therefore provides a basis for comparison. These statistics have also been used by a wide range of previous studies (Driver & Troutman 1989; Vaze & Chiew 2003).

Rainfall/runoff plots were used to assess the quality of runoff monitoring as well as to investigate the impact of pervious/impervious area runoff (Boyd et al., 1994; Shuster et al., 2005); a poor or inconsistent relationship between rainfall and runoff would indicate probable monitoring problems. Event pollution loads were calculated for all events. For discretely monitored events, these were calculated using the integration approach (the event hydrograph and measured discrete concentrations were used to calculate the EMC by dividing the event pollutant load by the event volume), while for event-composite monitored events, the load was calculated as the product of the measured EMC and event volume.

For each site simple statistical analysis was undertaken to determine site average EMC (μEMC), its standard deviation, and range. Using all measured EMC at one site, the SMC was determined as the sum of all measured event loads divided by the total volume of all measured events.

Event Mean Concentrations and Site Mean Concentrations - The results for μEMC and SMC were compared against the values previously reported in literature (Tables 1 and 2). The values were also compared among the sites and the effects of land-use were examined. The EMCs collected at one site were examined for cross-correlations between pollutants. In particular the correlation between EMCs of phosphorus and TSS concentrations which has been previously reported (Sartor et al. 1974; Xanthopoulos & Hahn 1993) was tested. The EMC data was also then correlated to the following event descriptors:

- Total event rainfall
- Total event volume
- Maximum rainfall intensity
- Average rainfall intensity
- Maximum flow rate

Maximum rainfall intensity, average rainfall intensity and flow rate have often been used to represent the energy input from the storm event (Desbordes & Servat 1984; Driver & Troutman 1989; Walden 1999), since a higher EMC may be expected in a more intense storm if maximum values are a good descriptor of the process.

As EMC data sets have often been shown to be log-normally distributed (Duncan 1999) the recorded EMCs were log-transformed and tested for normality using the using the Shapiro-Wilk distribution test, prior to correlation analysis. To examine the potential impact of one or two extreme data points having excessive effects on the correlation the Spearman non-parametric correlation coefficient (r_s) (Fowler et al. 1998) was also tested on the raw untransformed data.

Pollutant Loads - An attempt was made to examine whether a simple relationship exists between pollution loads and rainfall characteristics. There are considerable practical advantages in being able to predict pollutant loads using rainfall, rather than flow, since rainfall data is more widely available, and generally easier to measure. If a simple relationship between rainfall and pollutant loads can be derived, a practical stormwater pollution load model could be established, without the requirement any runoff monitoring. Whilst some authors have expressed the concern of ‘spurious self-correlation’ between pollutant loads and flow or rainfall (e.g. Kenney 1982), there is also a genuine risk of ‘throwing the baby out with the bath-water’, by ignoring useful relationships between one (easy to measure) variable and another (more difficult to measure) variable (Prairie and Bird 1989). Whilst pollutant load is a fundamental variable in its own right, rainfall is a much easier variable to measure directly than are pollutant loads and attempts to relate the two are both understandable and potentially useful for development and implementation of stormwater management strategies. The use of rainfall intensity also indicates that there are other causal factors beyond a mere volume – load relationship.

The approach of Duncan (1995), which was partially tested by Vaze and Chiew (2003) on a limited data set, was used. A similar approach has been previously tested for modelling of pollution concentrations, where an instantaneous rainfall intensity was related to a single concentration (Henderson & Moys 1987; Osborn & Payne 1990; Price & Mance 1978). Therefore, a power of the rainfall intensity was summed over the entire duration of an event and then correlated with the pollution load. In other words, the strength of the following relationship was tested:

$$Load = W \times A\sum_{i=1}^{n} I^{b} \qquad \textit{Equation 1}$$

where *Load* is the load of pollutant produced over the period of interest, I is rainfall intensity measured at short time periods (six minutes or less), A is the catchment area, n is the number of equal time intervals for which I was measured over the period of interest, W and b are calibration coefficients dependent on the catchment and pollutant characteristics, respectively. The term $\sum_{i=1}^{n} I^{b}$ is referred to as the rainfall function. The correlation between pollutant load and runoff rate (R) was also tested, as per previous studies (Grottker 1987; Huber & Dickinson 1988):

$$Load = C \times A\sum_{i=1}^{n} R^{d} \qquad \textit{Equation 2}$$

where C and d are calibration coefficients dependent on the catchment and pollutant characteristics.

The term $\sum_{i=1}^{n} R^{b}$ is referred to as the runoff function.

The important distinction in this approach, as opposed to many previous studies, is that the total event load is a function of the discrete rainfall (or runoff) intensity pattern during an event, rather than a single intensity parameter (e.g. maximum or average intensity) assumed to characterise the entire event.

Initial screening was carried out through regression analysis of the event load of TSS, TN and TP against each explanatory variable, for each of the sites, tabulating the coefficient of determination (R^2) for each combination. For the rainfall function ΣI^b and runoff function ΣR^d, the power coefficients *b* and *d* were optimised iteratively to achieve the highest possible R^2 at each site. Further analysis of the calibration parameters, the sensitivity and uncertainty of the approach will be undertaken in future research.

16 **Existence of First Flush** - Some stormwater management strategies are based on the supposition that there is a first flush of pollution at start of each event. Therefore, the existence of a first flush phenomenon was investigated using the large number of collected pollutographs of TSS, TN and TP. The cumulative pollution load (normalised by total event load) was plotted against the cumulative volume (normalised by the total event volume) for each event, consistent with the approach used in several other studies (Bertrand-Krajewski et al. 1998; Deletic 1998b; Gupta & Saul 1996; Ma et al. 2002). A first flush was defined if:

- 60% of the pollution load is carried by the first 20% of the flow.
- The cumulative pollution load curve – for the first 20% of flow - has a slope greater than 45°.

The first definition is used because it represents a situation where significant loads are carried early in the event and therefore can be related to the design of treatment measures. The second measure is used because it is easily demonstrated, and has traditionally been used as a measure of the phenomenon and is therefore widely understood (Bertrand-Krajewski et al. 1998).

17

18 Results and Discussion

19 *Event Mean Concentration and Site Mean Concentrations*

Pollution levels - Stormflow TSS EMC concentrations (Table 4) are significantly below those recorded in the 1980s and 90s (Table 1, and City Design 2003; GHD & EPA Vic 1981; RossRakesh et al. 1999; Sharpin 1993); on average TSS was around 80 mg/L, which is approximately one half of that reported in the

Fuchs' 2004 worldwide study and in the US NURP study, dating back 20 years (Tables 1 and 2). The results confirm the more recent findings by Smullen et. al (1999), who reported that on average TSS=78 mg/L in recent monitoring data (Table 2). The results suggest that there is a need to re-evaluate default values used in the design practice of stormwater treatment systems (e.g. in Australian practice the values shown in Table 1 are currently used for sizing and design (e.g. Wong 2005).

The lower TSS values may also be a result of refined monitoring procedures. Past reported studies in Melbourne and Brisbane (City Design 2003; RossRakesh et al. 1999), which recorded TSS EMCs similar to those reported in Fuchs et al. (2004), were derived from catchments which include sections of natural waterway. It is possible that these sections contribute TSS through in-stream processes (sediment scour and bank slumping, for example). This hypothesis is supported by the lack of similar differences in TN and TP concentrations (Tables 1-3). Another contributing factor may be the recorded reduction in atmospheric pollution, which is resulting in cleaner cities and therefore cleaner runoff. Substantial reductions of atmospheric concentrations of many trace elements have been reported over the period of interest (Lee et al. 1994; Var et al. 2000). Most importantly, the finer particles of health-related interest (PM_{10}) show reductions, both in emissions to air (Harrison 2002) and in concentrations in air (Wang et al. 2006). However, the larger suspended particulate matter shows much less change (Jakeman et al. 1992; Var et al. 2000). Improved air quality may contribute to reduced TSS concentrations in urban runoff, but it appears unlikely to be the sole cause.

SMC values are typically less than the simple site average EMC values, particularly for TSS, which suggests that large events may be contributing lower proportional loads. Land use, when industrial or residential catchments are considered, does not have a major influence on TSS, TP or TN event concentrations. Indeed Mt. Waverley, which is a light industrial catchment, has generally the lowest EMCs for all pollutants. The Richmond site has comparatively high pollutant concentrations, suggesting that the age of the catchment infrastructure may be influencing pollution behaviour in addition to hydrology. Two sites stand out: Narre Warren, which has comparatively high concentrations of TP and TN, and the Monash Roof, which has low concentrations of all three pollutants. Narre Warren is influenced by the prevalence of poorly-maintained septic systems in the catchment.

Stormflow TSS concentrations at all catchments are at least four times higher than in base flow (Table 4). For phosphorus the concentrations are similar in most cases, except for Narre Warren. Nitrogen is always higher in base flow, except in the case of Mt. Waverley (where it is similar in both storm and base flow) and significantly higher at Narre Warren (this small catchment would not normally have base flow but there are clear indications of inputs to the stormwater system from sewers and septics).

The composition of the μEMC of nutrients is presented in Table 6. Three of the sites have a relatively high percentage of TN in dissolved form – Narre Warren, the Monash Roof and Burwood East. This is expected at Narre Warren due to the septic influences. As expected the Roof site also had a high percentage of dissolved nitrogen due to the anticipated low concentrations of solids in roof runoff. Burwood East, however, is a typical suburban residential catchment. The high percentage of dissolved nitrogen corresponds to a high base flow concentration of TN, possibly indicating a contamination source in the catchment, or potential influences of interflow after storm events, resulting in leaching of soil nitrogen (Taylor et al. 2005).

In terms of metals far fewer events were captured, and therefore interpretation is more difficult. The two larger urban catchments Mt. Waverley and Richmond are comparable for most parameters, with the Roof site having significantly lower EMCs for all metals. Zinc concentrations are significantly higher than those reported elsewhere (Table 1) for the two larger sites, possibly because of the widespread use of zinc galvanised roofing products in Australia. Lead concentrations are intermediate between the published NURP results and the recently compiled US data sets (Table 2). Metals fractions are generally similar for the Gilby and Richmond catchments; the Roof site however had higher fractions of dissolved copper and zinc (51% compared to 27% and 20% for copper, 90% compared to 56% and 37% for zinc). This probably reflects the smaller amount of solids available for metals transport on the roof site.

Correlation between pollutants - There is usually some degree of correlation between different pollutants at a particular site. The degree of correlation is important because it allows for 'surrogate prediction' of the long-term concentrations of one pollutant, based on the concentrations of another. Such an approach can be used in modelling mean annual pollutant loads, amongst other things. TSS and TP are strongly correlated at most sites (r^2 between 0.55 and 0.89), the exception being Narre Warren, where TP concentrations are likely to be distorted by septic tank effluent. For all sites other than Narre Warren and Burwood East, the relationships between TP and TN are also relatively strong (r^2 between 0.37 and 0.94), while the TSS and TN relationships vary from weak to moderately strong (r^2=0.20-0.78).

Impact of hydrological characteristics – There are few strong relationships between EMC and hydrological characteristics for each site and pollutant. Maximum rainfall intensity and maximum flow rate are usually significantly correlated with TSS, however the correlations are not strong (r^2 ranging from 0.39 to 0.66). There is no clear pattern with TP and TN. Average rainfall intensity is significantly correlated on only 4 of the 19 cases (TSS at Narre Warren and Doncaster, TN at Narre Warren and TP at Burwood East) There is a significant negative correlation at Narre Warren between the total event rainfall and TP, and between total event rainfall and TN which is most likely caused by the previously discussed septic tank leakage into the drainage system; as the rainfall increases the runoff from septic systems is diluted causing a drop in concentrations.

Pollutant Loads

Table 7 shows the correlations between total event loads and the rainfall and runoff functions. The Monash Roof site is heavily influenced by two events with EMCs orders of magnitude above the average. When these two events are omitted the results follow a similar pattern to all other catchments (figures given in brackets). The most important point about these correlations is that the *rainfall function* outperforms the runoff and total rainfall measures in all cases except TP and TN at Narre Warren (affected by septic inputs), and TSS at Burwood East. Even these cases have r^2 values for the rainfall function of above 0.75. The b term in the rainfall function (and to a lesser extent the d term in the runoff function) is relatively consistent between pollutants, with average figures of 1.9, 1.5 and 1.2 for TSS, TP and TN respectively. Further work is required in terms of calibration and verification to provide more confidence around these figures. The higher value for TSS however is consistent with the hypothesis that TSS requires more energy to mobilise that TN which has a significant dissolved fraction. The rainfall function achieves a significantly greater correlation than the event volume in the great majority of cases, indicating that the estimation of loads could be greatly improved by using a rainfall-intensity based function. This demonstrates that the correlations achieved are not purely a result of a spurious correlation between runoff volume and load, as has been described in other load based models

(Huber 1986). These results support the hypothesis that a model based on rainfall intensity should perform better than a purely EMC based model (Chiew et al., 1997; Vaze & Chiew, 2003). Most importantly, this approach to modelling loads is simple and does not require flow modelling; i.e. after the relationship is established, a good load prediction is possible using only rainfall data without any flow monitoring.

20 First Flush

Although the pollutant load precedes the flow "load" on average in all cases, the effect is not strong (Table 8); the majority of events have loads preceding flows, however this is usually in the range of 25-30% of the total load being reported at the point where 20% of the volume has passed. For all pollutants the Monash Roof shows a slightly stronger effect that supports the explanation of first flush being a phenomenon of small catchments with simple drainage arrangements (Lee et al. 2002). Given that rainfall intensity is shown to be the significant driver of pollutant loads (Table 7), as opposed to flow volume, it is expected that loads would to some degree precede flow (as rainfall occurs before flow). This observation is in agreement with similar studies on two small asphalt sites in Europe (Bertrand-Krajewski et al. 1998; Deletic 1998b). Therefore, the management practices which target the first-flush (leaving subsequent flows untreated or treated to a much lesser extent) should be revised.

21 Practical Implications and Conclusions

The study described provides a unique data set of urban stormwater quality, collected using short timestep monitoring on a variety of urban catchments. Importantly, the monitoring programme was built on a thorough analysis of uncertainty and the number of samples needed to achieve the specified level of confidence in estimates of pollutant concentrations. Whilst the observed concentrations of TP and TN are consistent with earlier studies, the Melbourne sites showed significantly lower concentrations of TSS than in earlier monitoring programmes conducted in Australia and elsewhere. The lower concentrations support a similar observation of reductions in recent TSS concentrations, reported in the USA. This suggests a need to re-evaluate the concentrations used in the modelling and design of stormwater treatment systems.

Concentrations of nutrients, as well as most heavy metals were generally consistent with past studies. However, the EMCs of zinc were higher than those reported elsewhere.

Concentrations of TSS were higher during storms than during baseflow, whilst TN showed the opposite behaviour, appearing to be influenced by external (wastewater) impacts or by soil interflow processes during baseflow. The practical implication of this is that in many cases small events and baseflow cannot be bypassed, but must be treated.

Simple hydrological descriptors (such as total event rainfall, or event average rainfall intensity) were found to be relatively poor predictors of EMCs for all pollutants. However, a rainfall intensity function (based on an integration of rainfall, to a power, at each timestep) produced extremely good correlations with event loads, and is worthy of further investigation for stormwater pollutant modelling, as an alternative to simple EMC-based models. This approach could be particularly valuable for catchments without long-term flow monitoring.

A strong first flush was not observed at any of the sites, except for the smallest catchment (the Monash Roof), which has a very simple drainage layout, confirming observations from other recent studies. Therefore, management practices that treat only the first part of a storm, based on the presence of first-flush, should be revised.

22 5. References

Ahyerre, M., Chebbo, G., Tassin, B., & Gaume, E. (1998). Stormwater quality modelling, an ambitious objective? *Water Science Technology, 37*(1), 205-213.

Athayde, D.N., Shelley, P.E., Driscoll, E.D., Gaboury, D. & Boyd, G.

1983, Results of the Nationwide Urban Runoff Program, Environmental Protection Agency, Washington, DC., PB84-185537.

Ball, J. E., (2000), *Runoff from road surfaces - how contaminated is it?*, Proc. Hydro 2000 - Hydrology and Water Resources Symposium, Perth, WA, Australia, pp 259-264.

Behera P. K, Adams B. J, Li J.Y (2006). Runoff quality analysis of urban catchments with analytical probabilistic models *Journal of Water Resources Planning and Management-ASCE* 132 (1): 4-14 Jan-Feb 2006

Bertrand-Krajewski, J.-L. (in press). Influence of field data sets on calibration and verification of stormwater pollutant models. *Water Science and Technology*.

Bertrand-Krajewski, J. L., & Bardin, J. P. (2002b). *Uncertainties and representativity of measurements in stormwater storage tanks.* Paper presented at the 9th ICUD, Portland, Portland, Oregon, USA.

Bertrand-Krajewski, J. L., Barraud, S., & Bardin, J. P. (2002). Uncertainties, performance indicators and decision aid applied to stormwater facilities. *Urban Water, 4*, 163-179.

Bertrand-Krajewski, J. L., Chebbo, G., & Saget, A. (1998). Distribution of pollutant mass vs volume in stormwater discharges and the first flush phenomenon. *Water Research, 32*(8), 2341-2356.

Bertrand-Krajewski, J. L., Scrivener, O., & Briat, P. (1993). Sewer sediment production and transport modelling: A literature review. *Journal of Hydraulic Research, 31*(4), 435-460.

Boyd, M. J., Bufill, M. C., & Knee, R. M. (1994). Pervious and impervious runoff in urban catchments. *Hydrological Sciences Journal, 38*(6), 463-478.

Burton, G. A., & Pitt, R. E. (2002). *Stormwater effects handbook - A toolbox for watershed managers, scientists, and engineers* Lewis publishers.

Castaldi, F. J. (1983). *Final Report of the Nationwide Urban Runoff Program in Austin, Texas*. Austin: USEPA.

Chernick, M. R. (1999). *Bootstrap Methods: A Practitioner's Guide*. New York: Wiley.

Cordery, I. (1977). Quality Characteristics of Urban Storm Water in Sydney, Australia. Water Resources Research, 13 (1), 197 -201

City Design. (2003). *Stormwater Monitoring Program 2002-2003* (Report on Monitoring Program). Brisbane: City Design/Brisbane City Council.

Commonwealth of Australia. (2002). *The value of water: Inquiry into Australia's urban water management. Report of the Senate Environment, Communications, Information Technology and the Arts Reference Committee*. Canberra, Australia: The Parliament of the Commonwealth of Australia.

Deletic, A., & Orr, D. W. (2003). *Water quality of runoff from paved surfaces.* Paper presented at the IAHR Congress, Thessaloniki, Greece.

Deletic, A. B. (1998a). Evaluation of Water Quality Factors in Storm Runoff from Paved Areas. *Journal of Environmental Engineering, 124*(9), 869-879.

Deletic, A. B. (1998b). The first flush load of urban surface runoff. *Water Research, 32*(8), 2462-2470.

Desbordes, M., & Servat, E. (1984). *Solids in Urban Runoff: Statistical Analysis of French Experimental Data.* Paper presented at the 3rd International Conference on USD. 1984 Goteborg, Goteborg.

Deutsch, J. C., & Desbordes, M. (1981). *Study of Runoff Pollution for Urban Planning.* Paper presented at the Urban Stormwater Quality, Management, and Planning, Proceedings of the Second International Conference on Urban Storm Drainage, Urbana, Illinois. pp. 141-145.

Driver, N. E., & Troutman, B. M. (1989). Regression Models for Estimating Urban Storm-Runoff Quality and Quantity in the United States. *Journal of Hydrology, 109*, 221-236.

Duncan, H. P. (1995). *A Review of Urban Storm Water Quality Processes* (No. Report 95/9). Melbourne, Australia: Cooperative Research Centre for Catchment Hydrology.

Duncan, H. P. (1999). *Urban Stormwater Quality: A Statistical Overview* (No. Report 99/3). Melbourne, Australia: Cooperative Research Centre for Catchment Hydrology.

Fowler, J., Cohen, L., & Jarvis, P. (1998). *Practical Statistics for Field Biology* (2nd ed.). Chichester, England: John Wiley and Sons.

Francey, M., Deletic, A., Fletcher, T., & Duncan, H. P. (2004, 11-14 March 2004). *Monitoring of Stormwater Runoff in Melbourne.* Paper presented at the 19th European Junior Scientists Workshop - Process Data and Integrated Urban Water Modelling, Lyon, France.

Francey, M., Duncan, H. P., Deletic, A., & Fletcher, T. (2005). *Using a simple model to predict the behaviour of pollutant loads.* Paper presented at the 10th International Conference on Urban Drainage, Copenhagen, Denmark.

Fuchs, S., Brombach, H., & Weib, G. (2004). *New database on urban runoff pollution.* Paper presented at the Novatech, Lyon, France.

Greenberg, A. E., Clesceri, L. S., & Eaton, A. D. (1999). *Standard Methods for the Examination of Water and Wastewater, 20th Edition*. New York, USA: American Public Health Association, Water Environment Foundation, and American Water and Wastewater Association.

Grottker, M. (1987). Runoff Quality from a Street with Medium Traffic Loading. *The Science of the Total Environment, 59*, 457-466.

Gupta, K., & Saul, A. (1996). Specific Relationships for the first flush load in combined sewer flows. *Water Research, 30*(5), 1244-1252.

Gutteridge Haskins and Davey, & Environment Protection Authority of Victoria. (1981). *Characterisation of Pollution in Urban Stormwater Runoff* (No. AWRC Technical Paper No. 60). Canberra: Australian Water Resources Council.

Harrison, R. M. (2002). Key Pollutants - Airborne Particles. *Clean Air and Environmental Protection, 32*(4), 38-41.

Henderson, R. J., & Moys, G. D. (1987). *Development of a Sewer Flow Quality Model for the UK.* Paper presented at the 4th International Conference on Urban Drainage, Lousana, August 1987.

Huber, W. C. (1986). *Modelling Urban Runoff Quality: State of the Art.* Paper presented at the Urban Runoff Quality - Impact and Quality Enhancement Technology, Proceedings of an Engineering Foundation Conference, New Hampshire. pp. 34-48.

Huber, W. C., & Dickinson, R. E. (1988). *Storm Water Management Model, Version 4: Users' Manual*. Gainesville: University of Florida.

Jakeman, A. J., Taylor, J. A., & Symons, H. D. (1992). Efficient Tools for Analysising the Influence of Sources and Meteorology on Urban Ambient Concentration Trends Illustrated for Canberra, Australlia. *Ecological Modelling, 64*, 125-157.

Kenney, B.C. (1982). Beware of spurious self-correlations. *Water Resources Research*. 18(4), pp. 1041-1048.

Lee, D. S., Garland, J. A., & Fox, A. A. (1994). Atmospheric Concentrations of Trace Elements in Urban Areas of the Unitied Kingdom. *Atmos Environ, 28*(16), 2691-2713.

Lee, J. H., Bang, K. W., Ketchum, L. H., Choe, M. J., & Yu, M. J. (2002). First flush analysis of urban storm runoff. *The Science of the Total Environment*(293), 163-175.

Ma, J.-S., Khan, S., Li, Y.-X., Kim, L.-H., Ha, S., Lau, S.-L., et al. (2002). *First Flush Phenomena for Highways: How it can be meaningfully defined.* Paper presented at the 9th International Conference on Urban Drainage, Portland, USA.

Novotny, V. (2003). *Water Quality - Diffuse Pollution and Watershed Managment* (second ed.): John Wiley and Sons, inc.

Osborn, M. P., & Payne, J. A. (1990). *Calibration, Testing and Application of the Sewer Flow Model MOSQITO.* Paper presented at the 5th International Conference on Urban Storm Drainage, Osaka, Japan, July 1990.

Pitt, R., & Maestre, A. (2005, August 2005). *Stormwater quality as described in the National Stormwater Quality Database (NSQD).* Paper presented at the 10th ICUD, Copenhagen, Denmark.

Price, R. K., & Mance, G. (1978). *A Suspended Solids Model for Storm Water Runoff.* Paper presented at the Urban Storm Drainage, Proceedings of the International Conference held at the University of Southampton. pp. 546-555.

Prairie, Y.T. & Bird, D.F. (1989). Some misconceptions about the spurious correlation problem in the ecological literature. *Oecologia* 81, pp. 285-288.

RossRakesh, S., Gippel, C., Chiew, F. H. S., & Breen, P. (1999). *Blackburn Lake discharge and water quality monitoring program:Data summary and interpretation.* (Technical report No. 99/13). Melbourne: CRC for Catchment Hydrology/Freshwater Ecology.

Sartor, J. D., Boyd, G. B., & Agardy, F. J. (1974). Water Pollution Aspects of Street Surface Contaminants. *Journal of the Water Pollution Control Federation, 46*(3), 458-467.

Sharpin, M. G., (1993), *Hydrological variability, urbanisation and water quality in Australia*, Proc. 6th International Conference on Urban Storm Drainage, Niagara Falls, Canada, pp 543-548

Shuster, W. D., Bonta, J., Thurston, H., Warnemuende, E., & Smith, D. R. (2005). Impacts of impervious surface on watershed hydrology: A review. *Urban Water Journal, 2*(4), 263-275.

Smullen, J. T., Shallcross, A. L., & Cave, K. A. (1999). Updating the U.S. nationwide urban runoff quality database. *Water Science Technology, 39*(12), 9-16.

Taylor, G. D., Fletcher, T. D., Wong, T. H. F., Breen, P. F., & Duncan, H. P. (2005). Nitrogen composition in urban runoff—implications for stormwater management. *Water Research, 39*, 1982-1989.

Var, F., Narita, Y., & Tanaka, S. (2000). The Concentration, Trend and Seasonal Variation of Metals in the Atmosphere in 16 Japenese Cities shown by the Results of National Air Surveillance Network (NASN) from 1974 to 1996. *Atmos Environ, 34*, 2755-2770.

Vaze, J., & Chiew, F. H. S. (2002). Experimental study of pollutant accumulation on an urban road surface. *Urban Water, 4*(2002), 379-389.

Vaze, J., & Chiew, F. H. S. (2003). Comparative evaluation of urban storm water quality models. *Water Resources Research, 39*(10), 1280.

Vaze, J., & Chiew, F. H. S. (2003b). Study of pollutant washoff from small impervious experimental plots. *Water Resources Research, 39*(6), 1160.

Wang, X., Tong, S., Verall, K., Gerber, R., & Wolff, R. (2006). Air Pollution Trends in Brisbane, Australia, between 1980 and 2003. *Clean Air and Environmental Quality, 40*(1), 34-39.

Weiss P.T, Gulliver J.S, Erickson A.J (2007) Cost and pollutant removal of storm-water treatment practices *Journal of Water Resources Planning and Management-ASCE* 133 (3): 218-229 MAY-JUN 2007

Wong, T. H. F. (Ed.). (2005). *Australian runoff quality*. Sydney, Australia: Institution of Engineers, Australia.

Xanthopoulos, C., & Hahn, H. H. (1993). *Anthropogenic Pollutants Wash-off from Street Surfaces.* Paper presented at the Proceedings of the Sixth International Conference on Urban Storm Drainage, Niagara Falls, Ontario, Canada. Vol. 1, pp. 417-422.

Figure 1 - Location of Catchments

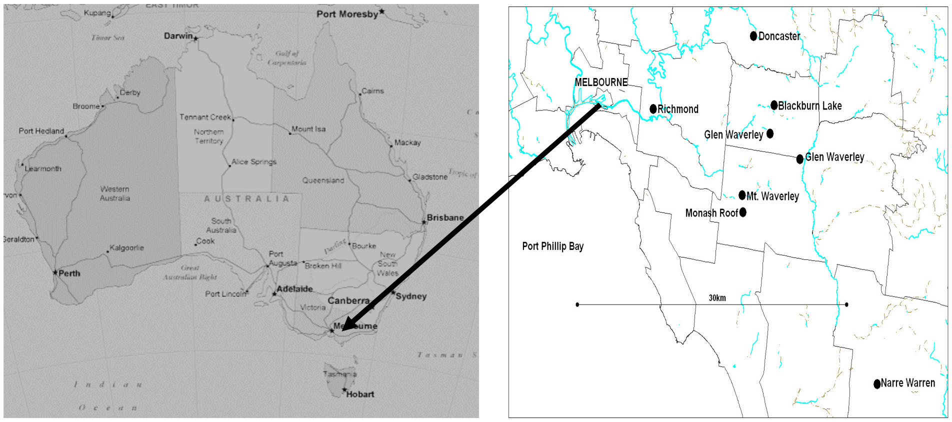


Table 1 – Total concentrations of the key pollutants adopted from a world wide compilation of urban runoff data set (Fuchs et al. 2004)

Parameter	Number of records	25th percentile	Median	75th percentile
TSS	178	74 mg/l	141 mg/l	279.7 mg/l
TP	149	0.24 mg/l	0.42 mg/l	0.70 mg/l
TN	17	2.1 mg/l	2.36 mg/l	5.8 mg/l
NH_4-N	41	0.5 mg/l	0.8 mg/l	1.20 mg/l
NO_3-N	110	0.60 mg/l	0.80 mg/l	1.49 mg/l
Cd	54	1.16 µg/l	2.3 µg/l	5.0 µg/l
Pb	158	46 µg/l	118 µg/l	239 µg/l
Cu	127	28 µg/l	48 µg/l	110 µg/l
Zn	128	128 µg/l	275 µg/l	502 µg/l

Table 2 – Comparison of compiled US data with NURP project data (adapted from Smullen et. al ,1999)

Data source	Parameter (Mean EMC)					
	TSS (mg/l)	TP (mg/l)	Nitrite/ Nitrate (mg/l)	Copper (µg/l)	Lead (µg/l)	Zinc (µg/l)
Compiled recent US Data	78.4	0.315	0.658	13.5	67.5	162
US NURP Data (>20 years old)	174	0.337	0.837	66.6	175	176

Table 3 – Summary of site details

Site Name *(Suburb)*	Primary land use	Area (ha)	Total fraction impervious
Mt. Waverley	Commercial	28.2	0.80
Glen Waverley	Medium Density Residential	38.0	0.45
Narre Warren	Rural Residential	10.5	0.20
Monash Roof	Treated Aluminium Roof	0.046	1.0
Burwood East	Mixed residential and commercial	186.0	0.46
Richmond	High Density Residential	89.1	0.74
Doncaster	Medium Density Residential	105.6	0.51

Table 4 – Summary of TSS, TP and TN concentration (σ = standard deviation)

Site	Par.	Wet Weather Events				Dry Weather			
		No of event	SMC (mg/l)	μEMC (mg/l)	σ of EMC (mg/l)	No of sample	Median (mg/l)	Mean (mg/l)	σ (mg/l)
Mt. Waverley	TSS	49	71.4	71.6	59.2	14	6.75	7.65	4.79
	TP	47	0.15	0.17	0.13	14	0.11	0.22	0.21
	TN	47	1.02	1.17	0.82	14	0.92	1.13	0.62
Glen Waverley	TSS	40	76.6	94.8	79.5	16	14.0	20.6	20.2
	TP	36	0.20	0.24	0.18	16	0.24	0.23	0.12
	TN	36	1.48	1.74	0.97	16	2.05	2.34	1.21
Narre Warren	TSS	41	79.6	91.9	65.8	16	8.45	10.0	6.01
	TP	17	0.66	0.75	0.20	16	7.95	9.01	4.45
	TN	17	3.56	3.51	0.72	16	31.5	32.6	9.52
Monash Roof	TSS	30	11.0	16.7	35.0	No dry weather flow			
	TP	30	0.02	0.03	0.04				
	TN	30	0.44	0.53	0.69				
Burwood East	TSS	32	54.0	84.1	93.2	16	5.55	7.27	5.99
	TP	16	0.16	0.15	0.06	16	0.37	0.63	0.84
	TN	16	1.58	1.54	0.54	16	2.25	3.41	3.70
Richmond	TSS	40	101.8	125.1	92.1	10	6.40	12.6	17.4
	TP	39	0.35	0.42	0.31	10	0.22	0.42	0.68

	TN	39	1.95	2.29	1.27	10	9.55	11.6	10.6
Doncaster	TSS	52	78.1	77.0	56.3	10	13.0	16.0	10.3
	TP	No event monitoring of TP or TN				10	0.22	0.24	0.13
	TN					10	2.05	2.39	1.09

Table 5 – Summary of heavy metals and other pollutant data

	Roof			Mt. Waverley			Richmond		
Parameter	Number of events	μEMC	σ of EMC	Number of events	μEMC	σ of EMC	Number of events	μEMC	σ of EMC
pH	17	5.7	0.42	18	6.6	0.25	13	6.7	0.24
EC (uS/cm)	17	17.2	16.80	18	57.3	17.41	13	102.6	37.56
alkalinity (mg/l)	17	3.5	2.82	18	14	4.45	13	26.6	7.95
Ca (mg/l)	17	0.8	0.92	16	4.8	1.07	13	8.2	1.90
Mg (mg/l)	17	0.2	0.22	16	0.9	0.28	13	1.4	0.50
Na (mg/l)	17	1.8	1.76	16	4.9	1.25	13	7.2	2.70
K (mg/l)	17	0.2	0.12	16	1.1	0.54	13	2.5	0.98
Cl (mg/l)	17	2.4	3.56	16	5.1	2.40	13	8.9	5.85
SO_4 (mg/l)	17	1.5	1.00	16	4.5	1.97	13	6.4	2.23
TOC (mg/l-C)	17	3.1	1.81	16	8.2	4.21	13	11.7	5.81
DOC (mg/l-C)	14	2.4	0.94	16	5.3	2.53	12	6.9	2.78
Al (mg/l)	12	0.89	1.17	11	7.98	14.76	8	5.30	3.15
Cd (μg/l)				10	1.3	2.6	8	0.5	0.3
Cr (μg/l)	16	3.2	4.8	18	17.3	32.2	13	13.2	6.6
Cu (μg/l)	17	6.3	6.2	18	46.3	60.7	13	56.1	25.5
Fe (mg/l)	17	1.12	1.95	18	7.40	15.10	13	6.70	3.90
Mn (μg/l)	17	21.0	26.6	18	113.5	236.8	13	116.9	54.7
Ni (μg/l)	12	2.9	3.2	11	24.0	44.3	8	16.8	6.5
Pb (μg/l)	15	5.8	6.0	18	43.3	78.4	13	112.2	60.8
Zn (mg/l)	17	0.12	0.07	18	0.95	0.79	13	1.24	0.43

Table 6 – Speciation of nitrogen and phosphorus as percentage of EMCs

	% of TN that is					% of TP that is
	NH_3	NOx	Total Dissolved Nitrogen	Dissolved Organic Nitrogen	Particulate Organic Nitrogen	Filterable Reactive Phosphorus
Mt. Waverley	4%	29%	59%	25%	41%	12%
Glen Waverley	6%	32%	58%	20%	42%	25%
Narre Warren	21%	41%	79%	17%	21%	58%
Monash Roof	27%	27%	71%	18%	29%	25%
Burwood East	5%	51%	77%	21%	23%	37%
Richmond	8%	28%	57%	20%	43%	20%

Table 7 – Correlations between event loads and rainfall and runoff functions (the best explanatory variable in each case is denoted by bold).

Catchment	Pollutant	Total event rainfall, ΣI	Rainfall function, ΣI^b	Total event runoff, ΣR	Runoff function, ΣR^d
Mt. Waverley	TSS	0.78	**0.91**	0.68	0.88
	TP	0.88	**0.89**	0.87	**0.89**
	TN	0.92	**0.94**	0.85	0.91
Glen Waverley	TSS	**0.83**	**0.83**	0.65	0.73
	TP	**0.81**	**0.81**	0.67	0.73
	TN	**0.89**	**0.89**	0.77	0.85
Narre Warren	TSS	0.88	**0.91**	0.82	0.87
	TP	0.87	0.89	**0.93**	**0.93**
	TN	0.94	0.94	**0.98**	**0.98**
Monash Roof*	TSS	0.60 (0.67)	**0.62 (0.93)**	0.38 (0.67)	0.46 (0.91)
	TP	0.59 (0.82)	**0.60 (0.85)**	0.56 (0.78)	0.58 (0.81)
	TN	**0.76 (0.88)**	**0.76 (0.88)**	0.74 (0.85)	0.75 (0.87)
Richmond	TSS	0.73	**0.90**	0.68	0.73
	TP	0.82	**0.90**	0.79	0.79
	TN	0.91	**0.94**	0.89	0.89
Burwood East	TSS	0.93	0.93	**0.94**	**0.94**
Ruffey’s Wetland	TSS	0.94	**0.96**	0.93	0.93

* The Monash Roof site was heavily influenced by two events with EMCs several orders of magnitude higher than the mean. Figures in brackets show results with these two outliers excluded.

Table 8 – First Flush Analysis

Site/Criteria	Pollutant	Average % Load carried by the first 20% of the flow	% of Events with 60% of load carried by the first 20% of volume
Mt. Waverley	TSS	27	4
	TP	25	0
	TN	24	0
Glen Waverley	TSS	24	0
	TP	23	0
	TN	22	0
Narre Warren	TSS	29	3
Monash Roof	TSS	39	4
	TP	45	2
	TN	36	0
Burwood East	TSS	27	0
Richmond	TSS	28	0
	TP	28	0
	TN	27	0
Doncaster	TSS	29	5

Appendix F - Paper under review in Journal of Environmental Engineering (Submitted April 2008, Revised copy submitted October 2009)

The testing and sensitivity of a simple method for predicting urban pollutant loads

AUTHOR NAMES: *M. Francey*, H.P. Duncan, A. Deletic, T.D. Fletcher*

AUTHOR ADDRESS: **Melbourne Water, East Melbourne 3001, Australia* (e-mail: matt.francey@melbournewater.com.au)

AUTHOR FOOTNOTES:

Matt Francey Melbourne Water, 100 Wellington Parade, East Melbourne, VIC, AUSTRALIA, 3001. Tel: +61 3 9235 2221 Fax: +61 3 9235 7200

Dr Timothy Fletcher, Director, Institute of Sustainable Water Resources, Department of Civil Engineering, Faculty of Engineering, Monash University, Tel +61 3 990 52599; tim.fletcher@eng.monash.edu.au

Associate Professor Ana Deletic, Institute of Sustainable Water Resources, Department of Civil Engineering, Faculty of Engineering, Monash University, +61 3 990 52940; ana.deletic@eng.monash.edu.au

Hugh Duncan, Project Researcher , Institute of Sustainable Water Resources, Department of Civil Engineering, Faculty of Engineering, Monash University, +61 3 990 54981; hugh.duncan@eng.monash.edu.au

ABSTRACT

This paper tests a simple two-parameter regression model, based on rainfall intensity, for calculating event loads of total suspended solids (TSS), total phosphorus (TP) and total nitrogen (TN) from urban catchments. It also examines the sensitivity of the model to its two parameters, and to the rainfall time step. This was done using large data sets collected at six urban catchments in temperate Melbourne. It was found that the two-parameter model typically explains about 90% of the variation in event loads at a site.

The model also predicts the within-event behaviour of pollutants when the flow lag time is taken into account, with R^2 correlations greater than 0.6 in most cases for both loads and concentrations at a six minute timestep. Despite its acknowledged correlation with flow, rainfall intensity over short timesteps is shown to be the main driver of pollutant mobilisation, and provides a practical means of predicting pollutant loads using readily available data.

KEYWORDS: Urban Stormwater, Water Quality, Monitoring, Total Solids, Nitrogen, Phosphorus, Rainfall Intensity

Introduction

The widespread use of stormwater quality models in urban water management has been hindered by their unreliability (Bertrand-Krajewski, Scrivener et al. 1993; Ahyerre, Chebbo et al. 1998). particularly in catchments where little flow and water quality data available (the case in most practical model applications). Models currently in use range from statistical approaches through to increasingly complex process-based models, as briefly outlined below.

Statistical and *Regression models* use monitoring data to relate concentrations or loads to hydrological or catchment characteristics. They range from simple models that typically multiply the event-mean concentration (EMC) by the event runoff volume to determine the pollutant load to probabilistic models that use more sophisticated approaches (Rossi et. al. 2005). The difference in EMCs between sites however can be substantial, even for catchments with similar land use (Chiew and McMahon 1998), so the use of statistical models, without calibration to a particular catchment, involves considerable uncertainty.

Multiple regression approaches where hydrological or catchment factors have been linked to pollution behaviour have been used by many authors (Bradford 1977; Jewell and Adrian 1982; Tasker and Driver 1988; Driver and Troutman 1989). Desbordes and Servat (1984) produced two equations based on the data from the four catchments to formulate models for TSS, BOD and COD loads. Maximum 5-minute rainfall intensity and runoff volume during a storm were used as variables initially and surface loading was added in further trials. Both equations performed similarly, with errors of up to 10% for annual totals and 30% for individual events.

Models such as these can produce good medium to long-term predictions of pollutant loads when calibrated, and are therefore appropriate for uses such as the prediction of annual pollutant loads.

Physically based models attempt to relate not only the total pollutant loads generated but the behaviour within an event by attempting to replicate the dominant processes involved. Models are usually built on the paired concepts of buildup and washoff. Buildup is the process by which dry deposition accumulates on impervious areas (Sartor, Boyd et al. 1974; Pitt 1979; Novotny, Sung et al. 1985; Bertrand-Krajewski, Scrivener et al. 1993). Washoff is the process by which accumulated dry deposition is removed from impervious surfaces by rainfall and runoff and is incorporated in the flow. Washoff loads have been estimated using variables such as total runoff volume (Characklis, Bedient et al. 1978; Mance and Harman 1978), total event rainfall (Reinertsen 1981), runoff rate (Meyer 1969; Aalderink, Van Duin et al. 1990), rainfall intensity (Yaziz, Gunting et al. 1989; Kuo, Yan et al. 1993; Brezonik and Stadelmann 2002) or a combination of these factors. All four explanatory variables are to some extent correlated with each other, more so at longer time intervals, so it is difficult to tell which is the underlying driving force behind the observed washoff behaviour. Explanations of pollutant mobilisation have centred around the processes of impact energy from rainfall drops (Price and Mance 1978; Coleman 1993) or the shear stress from overland and channel flow (Akan 1987; Deletic, Maksimovic et al. 1997). Recently the complexity of more complex stormwater models has been questioned, with some studies showing that the number of parameters, rather that the physical representiveness, being responsible for calibration success (Ahyerre et. al. 1998, Kanso et. al. 2004).

This paper focuses on simple regression modelling approached, because of their implicit simplicity. Duncan (1995), in a review of urban stormwater processes, noted the various methods of representing rainfall intensity in stormwater quality modelling, proposing an equation which related event loads to rainfall intensity (Equation1)

$$eventload = \sum_{i=1}^{n} (aI^{b}) \qquad \text{(Equation 1)}$$

where I is the rainfall intensity as recorded in each of the *n* time steps, and a and b are calibration coefficients. It should be noted that *I* is calculated by assuming that total rainfall depth, recorded in one time step, occurred just within that very time step. Therefore, *I* strongly depend on the resolution of rainfall records.

The cumulative nature of this equation represents the ongoing input of energy through raindrop impact. The erosive effects of increasing rainfall intensity, and more particularly the power of the rainfall intensity, are often used in studies of pervious area runoff (see, for example Hudson 1971). Similar relationships have been shown to be appropriate for urban catchments (Vaze and Chiew, 2003b).

Vaze and Chiew (2003) used this approach on a modest data set to compare pollutant load predictions from Equation 1 at 6 minute timesteps, with alternative formulations based upon total event rainfall, total event runoff, sum of the runoff rate (substituting flow volume for intensity in equation 1), and a combination of the rate and intensity terms. They also compared each of these results with a typical buildup/washoff based model. The results for the Σrainfall intensity and Σrunoff rate models taken separately were similar (within 20% of recorded loads). When intensity and runoff functions were combined they generally achieved a higher correlation, but this may be partly due to having two extra calibration coefficients and only a small number of events (Ginzburg and Jensen 2004). The intensity and runoff functions performed as well or better than the process based (buildup and washoff) model.

The Σrainfall intensity and Σrunoff rate models was tested further on the data set presented in this paper (XXXXXX). The sum of rainfall intensity showed higher correlations than the Σrunoff rate with TSS, TP and TN in all cases except one (Burwood East with a R^2 of 0.94 to 0.93 for Σrainfall intensity).

Other authors have also been testing these concepts. Recently, Brodie (2007) tested very similar model (based on the square of the rainfall intensity at 6-minute timesteps) on small (<450m^2) catchments and recorded moderate agreement between the modelled and measured non-course particulate loads. The study further developed a “rainfall detachment index” based on rainfall intensity over short timesteps and the peak rainfall intensity. The index showed good agreement for the majority (but not all) of storms on these small catchments.

It could be concluded that more research is needed to gain confidence in the proposed simple regression model (Equation 1). An approach based on rainfall intensity, representing the kinetic energy contained in rainfall, seems to reflect at least TSS loads better than approaches based on runoff. If this approach is to be used on non-monitored catchments (the ultimate test), we need to understand the calibration parameters and if possible assess their values for different site conditions. These were the key objectives of the research undertaken by the authors.

This paper reports on attempts to test the proposed regression model on 6 catchments in Melbourne, Australia, for which very good quality data sets were available. The paper examines the behaviour of the model's two parameters, the sensitivity of the model to them, as well as their transferability between the catchments. Finally, it examines the ability of the model to predict pollution flux (concentrations and loads over short time steps) during individual storm events. In this work it is fully recognised that an applied model such as the proposed one is a balance between complexity and practical application.

Methods

Model formulation

The model is formulated as,

$$Load = W \times A \times T \times \sum_{i=1}^{n} I^{b} \qquad \text{(Equation 2)}$$

where *Load* is the load of pollutant produced by the whole catchment over the period of interest (units of mass), *W* is a calibration coefficient dependent on pollutant and location, *A* is the total catchment area (units of length^2), *T* is the duration of the timestep (units of time), *I* is the rainfall intensity in the timestep (units of depth/time), *b* is a constant dependant on the pollutant, and *n* is the number of time steps over the period of interest. The right hand side of the equation is referred to as the rainfall function.

The method described was primarily developed to predict the load over an event. By applying Equation 2 at short timesteps however it was also used to to predict pollution flux (i.e. load or concentration) within events. Whilst the model described here is not coupled with a hydrological model and therefore does not implicitly consider concepts such as lag time or the attenuation of flow, the influences of these behaviours on the model were be investigated.

Data for model testing

Six catchments in the south-eastern suburbs of Melbourne are used for the analysis. All catchments have separate piped stormwater systems, and were selected to cover a range of sizes and land uses (Table 1). Catchment areas and indicative total fraction impervious were calculated from GIS data and checked against aerial orthophotos. Melbourne has a temperate climate, with warm summers and cool winters. Rainfall occurs throughout the year, with a slight spring maximum. Mean annual rainfall across the area of interest ranges from 600 to 800 millimetres per year. At each catchment, a 0.2mm tipping bucket rain-gauge was installed as close as possible to the catchment centroid. Flow gauges were installed within the stormwater pipes at the outlet of each catchment, and used to trigger autosamplers, at pre-determined intervals based on estimated catchment flow and time of concentration. Further details regarding the collection of runoff, rainfall and pollutant data are provided in Francey *et al.* (2005).

Table 1 – Catchment and event descriptions

Site	Primary Land Use	Area (ha)	Fraction Impervious	No. of Events (TSS,TP,TN)	Comments on Monitoring
Mt. Waverley	Commercial	28.2	0.80	49, 48, 48	24 TP and TN from flow-weighted composites
Glen Waverley	Medium Density Residential	38.0	0.45	19, 17, 17	12 TP and TN from flow-weighted composites
Narre Warren Sth	Rural Residential	10.5	0.20	38, 17, 17	All TP and TN from flow-weighted composites
Burwood East	Mixed	186.0	0.46	29, -, -	10 TSS events with one -minute rainfall data, 29 with 6- minute
Richmond	High Density Residential	89.1	0.74	40, 39, 39	18 TP and TN using flow-weighted composites
Doncaster	Medium Density Residential	105.6	0.51	52	No TP or TN monitoring

Model calibration, verification, and transferability

To develop the model an "event based" approach is used. For the rainfall function ΣI^b the power coefficient *b* is optimised iteratively to achieve the highest possible R^2 at each site. The cross calibration approach of Vaze and Chiew (2003) was then used to calibrate the model. One event was left out of the dataset and calibration parameters were calculated. These parameters were then used to obtain a result for the event in question.

Initially the method was used to test the method for each individual catchment and pollutant. Based on the results of this the method was then applied to catchments not used to derive the parameter set. In this way the transferability of model parameters from one site to another was examined.

The calibration parameters were then used for a verification exercise and the results measured using three indicators:

- The ratio of the observed versus predicted loads for the total dataset.
- The co-efficient of determination of the observed versus predicted loads. The R^2 value compares the magnitude of the predicted versus observed loads for each event to other events in the data set but will not describe any systematic bias.
- The co-efficient of efficiency (E) of the observed versus predicted loads (Nash and Sutcliffe 1970). The E value compares the predicted results to the 1:1 (predicted:observed) line on a graph, therefore picking up any systematic bias in the results. The E value is always less than or equal to a corresponding coefficient of determination value, an E value of 0 indicates that the mean observed value is an equivalent predictor to the parameter tested (Legates and McCabe 1999).

The model parameter sensitivity

Given the calibration/verification exercise described above the model was then tested for the sensitivity to these calibration parameters. In an ideal situation these parameters would represent completely different physical processes or situations, however many models have been shown to have parameters that are dependent on one another – therefore leaving open the possibility of various solution sets (Kanso, Gromaire et al. 2003).

To evaluate model sensitivity towards its parameters, the model reliability (i.e coefficient of efficiency E) was estimated for a large range of the values that these parameters can take. For instance a fixed value of b was used and then W was trialled over a wide range and E was calculated for each parameter combination; b was then increased and the process repeated. To interpret the results each b/W pair which achieved a E value of greater than 0.6 was plotted with calibration parameters W and b on the x and y axis respectively. The shape of the objective function graphs tells us a lot about parameters and the model identifiability (if independent, the parameters would be represented on these graphs as a ”cloud”, demonstrating a clear optimal value for each parameter, while a linear plot would indicate multiple solution sets, and therefore a lack of identifiability).

Secondly, the effect of timestep was examined by testing the correlation between predicted and observed loads for 1, 6, 12 and 30-minute timesteps. The timestep at which rainfall is measured affects the relationship in two distinct ways. Firstly there is the direct effect on timestep load, accounted for by the factor *T* in Equation 2. There is also another effect, which is less easily accounted for, as it involves a real loss of information. If load is assumed to be proportional to a power of rainfall intensity which is greater than one the calculated load is generated disproportionately from the short periods of highest rainfall intensity. If a longer timestep is used the influence of short periods of high intensity rain is reduced, so $AT\Sigma I^b$ will be lower, and the fitted value of *W* will be correspondingly higher. The size of this effect depends upon the actual rainfall intensity pattern of a given event. Hence *W* depends on the timestep, in a way which cannot easily be generalised. At a practical level, the solution is to compare sites at the same timestep length. It is hypothesised that there will be an optimum relationship between the model timestep, size of rainfall tip and the catchment size. If the timestep is too small the model will take on a “binary” process, with each timestep either representing one tip of the gauge or no information. The effect of different timesteps will be assessed by the change in R^2 between the rainfall function and event load.

Thirdly, the parameter sets obtained for different catchments and pollutants were compared. The power or exponent, *b,* controls the shape or curvature of the relationship between rainfall intensity and pollutant load, and depends on the water quality parameter being modelled. The power *b* is dimensionless, but the washoff coefficient *W* has the difficult units of mass × $\text{length}^{-2} \times \text{time}^{-1} \times (\text{length/time})^{-b}$. When $b = 1$ this simplifies to units of concentration, but in the general case the units of *W* depend upon the power *b*. The practical significance of this is that the washoff coefficient *W* can only be compared between catchments when *b* is consistent between catchments. This was undertaken by taking an optimum timestep, then running the model using an optimum “b” value. The calibration/verification process used was repeated by removing each catchment form the calibration process in turn. The “W” value for all events was then presented as a distribution.

Modelling pollution flux

The aim is to examine whether pollution flux (load over very short time step) could be predicted using the proposed approach. However, right at the start it had to be acknowledged that the proposed approach does not account for the time-lag and attenuation of pollutants that are observed in practice; although generated by rainfall, both water and pollution travel along catchments and therefore their response lags, and is “smoothed by attenuation, after the rainfall. (Huber and Dickinson 1988). To test the potential for incorporating the proposed model with a hydrological model capable of accounting for lag and attenuation of flows, the impact of time lag and attenuation had to be assessed.

To study both time-lag and attenuation the following procedure was used. For each individual event, the rainfall versus observed flow at each timestep was compared for a lag time ranging from zero (when the rainfall is taken at exactly the same timestep as the flow is measured at the outlet) to 65 minutes. The latter (65 minutes) is beyond the time of concentration of any of the catchments using traditional methods such as those described in Australian Rainfall and Runoff (1997). Attenuation was simulated by taking a running average of the predicted load or concentration over the chosen attenuation time.

R^2 values for observed and predicted loads (and concentrations) were then calculated for attenuation times of 6, 10, 20 and 40 minutes and three different lag times - no lag, the optimum for each event and the average lag for the catchment (based on the average of all optimum lags). The analysis was conducted for both load (over a six minute timestep) and concentration (by dividing the predicted flow rate of pollutants by the observed flow).

Many physical models assume that there is a reduction in pollutant washoff with time due to less pollution being available on the surface later in a storm (Huber and Dickinson 1988). An analysis was undertaken to assess if there was any decline in pollutant loads relative to rainfall intensity as the event proceeds, and therefore whether there is a need for a depletion term in the model. This was undertaken by determining whether there is a decline in “W” values with time and event volume. For every point in time for which a water quality sample was taken a “W” value was calculated, by dividing the observed load for the timestep (or concentration at the time of measurement) by the predicted load/concentration.

The cumulative storm volume and time elapsed, for the same point in time in the event, was recorded. The parameter W was than plotted against the storm volume and time elapsed in the event. Any indication of depletion should be indicated by a trend in a graph of “W” values, therefore the slope of a linear trendline was assessed for its significance (at a 0.05 level).

Results and Discussion

The results are presented in terms of the event load analysis, then the model sensitivity and finally the within event analysis.

Model calibration, verification, and transferability

When the calibration exercise was undertaken the b parameters calculated (for a one-minute timestep) optimises on distinct values for different pollutants; 1.8 for TSS, 1.4 for TP and 1.2 for TN. From these optimised values a set of calibration parameters (W) were produced that will allow comparison with other sites. The mean W values (and coefficient of variation) were 0.76 (41%), 0.0019 (54%) and 0.0011 (29%) for TSS, TP and TN respectively.

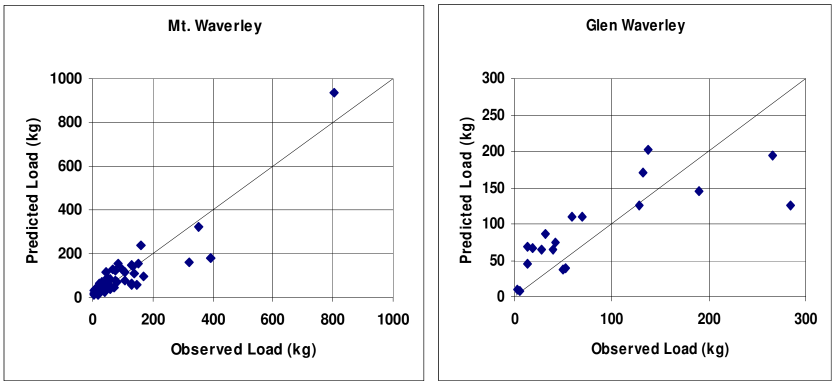
The verification exercise demonstrated that when each catchment is considered individually the total loads at each site are predicted well (see Figure 1 and Table 2), with a maximum percentage error of 13% (TSS at Richmond), the average error being 7%, 5 % and 6% for TSS, TP and TN respectively.

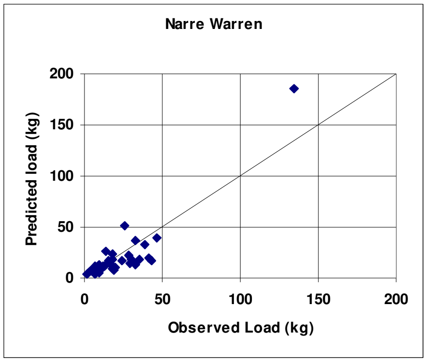
When the prediction of individual events is considered the method achieves coefficient of efficiency values typically above 0.5 and usually above 0.7 for all pollutants. The only exception to this is TP for the Narre Warren, which returns a value of – 0.04 (remembering that a negative value indicates that the mean load is a better predictor than the function tested). This is largely due to one event in the dataset which was estimated as being in excess of 1 in 50-year recurrence interval for the area and caused significant flooding. The Narre Warren catchment has a significantly greater pervious proportion than other catchments as well as being influenced by septic systems (Francey, 2005), and it is likely that this result represents a combination of flood-related pollutant transport behaviour.

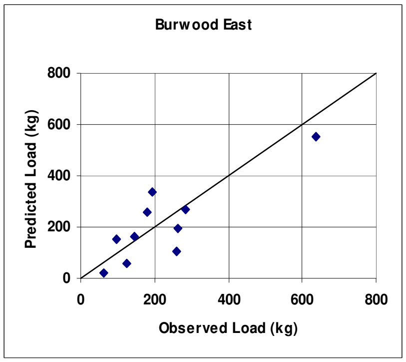
Table 2 – Verification of individual catchments for event loads

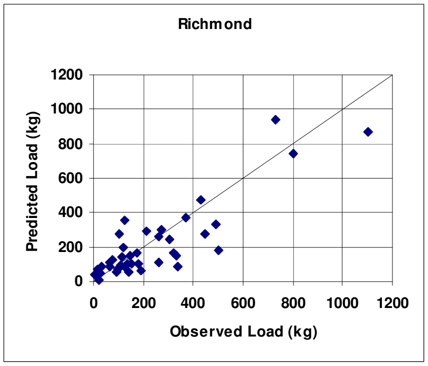
Catchment	Pollutant	Observed/Predicted Total Load	R^2 of Observed vs Predicted events	E of Observed vs Predicted events
Mt. Waverley	TSS TP TN	0.99 1.00 1.01	0.84 0.71 0.88	0.82 0.65 0.88
Glen Waverley	TSS TP TN	0.89 0.93 0.96	0.61 0.47 0.67	0.58 0.46 0.67
Narre Warren Sth	TSS TP TN	1.08 1.03 0.95	0.83 0.79 0.85	0.67 -0.04 0.56
Burwood East	TSS	1.07	0.71	0.69
Richmond	TSS TP TN	1.13 1.11 1.06	0.76 0.70 0.79	0.74 0.68 0.79
Doncaster	TSS	0.92	0.94	0.92

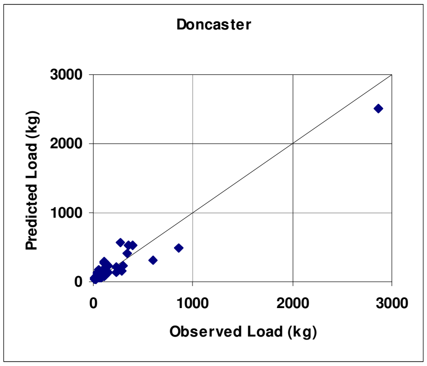
Figure 1 – Observed vs predicted loads (TSS)









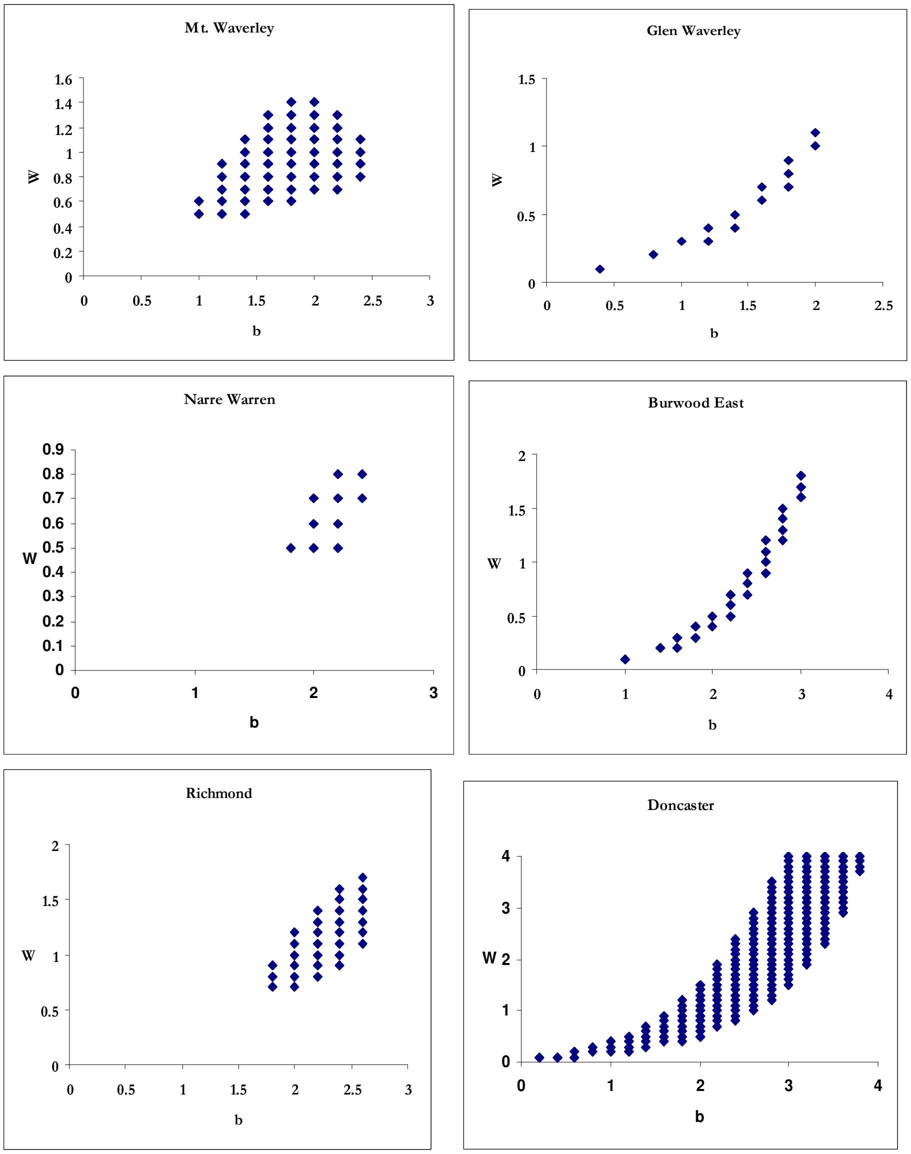


When the same exercise was repeated using cross calibration at a catchment level the results are also encouraging. In terms of total loads the average differences between observed and predicted are 23, 28 and 18% for TSS, TP and TN respectively. All E values for individual events are above 0.45 except for TP at Glen Waverley and TSS and TP at Burwood East. Essentially these predictions are for completely independent catchments as none of the specific catchment data tested has been used to calibrate the model.

The proposed model is actively seeking simplicity and attempts to represent the most critical processes for pollutant washoff. Indeed every extra detail and every extra parameter to be fitted must be paid for with more data to achieve a satisfactory fit, and with more noise introduced along with the additional data. See Ginsberg and Jensen (2004) for a discussion of these concepts. More detail can be added if necessary as future data demands and permits. Hence, whilst it is true that pollutant loads intrinsically incorporate an element of rainfall in them, the use of rainfall (rather than flow) provides extra descriptive ability to the model. However, whilst this concept of 'spurious self-correlation' has been noted in the literature (e.g. Kenney 1982), there is also a genuine risk of 'throwing the baby out with the bath-water', by ignoring useful relationships between one (easy to measure) variable and another (more difficult to measure) variable (Prairie and Bird 1989). Whilst pollutant load is a fundamental variable in its own right, flow is a much easier variable to measure directly than are pollutant loads and attempts to relate the two are both understandable and potentially useful for development and implementation of stormwater management strategies. In this sense, it is critical to distinguish between 'spurious correlation' and 'spurious inference' (Prairie and Bird 1989). We do not claim that the high correlations between the rainfall function and pollutant loads are a function of rainfall energy alone; but the model does account for a high degree of the variability in pollutant loads, because it describes both the flow (on which loads obviously depend) and in accounts for the *additional influence* brought by rainfall energy. In summary, using the rainfall function provides explanatory power, using an easy to measure variable, and has low cost in terms of calibration when compared to complex models.

The model parameter sensitivity

The model efficiency was plotted against combinations of the model parameters when E > 0.6. There were no cases of multiple solutions in any of the graphs; however some of the graphs have highly linear plots, indicating solution sets over a range of *b* values.   Most graphs centred on a peak, however some showed a relative insensitivity to the parameter *b* which was reflected in a "ridge" where high E values where achieved across a wide variety of values (Figure 2). The wide range of shapes is likely to be a representation of the representiveness of the data sets, with catchments such as Mt. Waverley and Richmond, which had a wide range of event sizes in the data set, showing a less linear response than Doncaster, which is dominated by one large event.


Mt. Waverley

Glen Waverley

Narre Warren

Burwood East

Richmond

Doncaster

Figure 2 - Plot of W/b combinations (TSS) correlation with observed load (E > 0.6)

When the effect of timestep was considered it was expected that the 1-minute would give superior results to longer timesteps. This was generally the case however the 6 and 12-minute timesteps also give high correlations (average R^2 values of 0.82, 0.79, 0.77 and 0.69 for the 1, 6, 12 and 30 minute timesteps, respectively). Given the generally relatively high correlations for a 6-minute timestep, and the widespread availability of 6-minute rainfall data in Australia, the decision was made to proceed with a 6-minute timestep for further model development. The declining correlations with timestep suggest why these relationships have been less evident in past studies (Brezonik and Stadelmann 2002). When long rainfall timesteps are used it has the impact of smoothing the power function associated with rainfall intensity.

The calibration/verification exercise was repeated using six minute data and it produced results similar to the "independent catchment" exercise in section 3.1. As discussed in the sensitivity analysis the model is shown to be relatively insensitive to the parameter b for some catchments. Encouragingly however, the b parameter however optimises on distinct values for different pollutants; 1.5 for TSS, 1.3 for TP and 1.0 for TN. From these optimised values a set of calibration parameters (W) were produced that will allow comparison with other sites. The mean W values (and coefficient of variation) were 0.35 (66%), 0.0013 (65%) and 0.0063 (59%) for TSS, TP and TN respectively.

These results accord with the role of the relative importance of energy input, in that the importance of rainfall energy increases with the proportion of a pollutant that is in particulate form (and consequently *b* is higher for particulate than dissolved pollutants). This observation supports previous studies on similar models which have found the power terms are pollutant-specific (Sriananthakumar and Codner 1992). Interestingly, nitrogen has an average b value of 1 which indicates that the total rainfall during the event explains nitrogen behaviour. This accords with the findings of Taylor *et al.* (2005), who found that 75% of nitrogen in stormwater is in the dissolved form. Other studies have found that also nitrogen concentrations in rainfall are similar to that in stormwater (Randall et al. 1981; Miller and Mattraw 1982), reinforcing the role that rainfall quantity, rather than energy *per se*, has in determining loads of nitrogen in stormwater.

Modelling pollution flux

When testing the use of average , optimum and no-lag times on the prediction of pollutant flux in all cases the use of optimum lag time significantly improves the performance of the prediction over the catchment average lag, or when using no lag (with average correlations of 0.20 for no lag, 0.35 for the average lag and 0.64 for all catchments). The model is not as sensitive to attenuation time, with the highest correlations being obtained when the attenuation is approximately the same as the average lag for the catchment. When the optimum lag is used the correlations for both loads over a 6-minute timestep and concentration are generally greater than 0.6 (Table 2). Although less events were available for TP and TN the same pattern is evident; R^2 (load) correlations for TP of 0.64, 0.79 and 0.66 and 0.70, 0.53 AND 0.65 for TN for Mt. Waverley, Richmond and Glen Waverley respectively.

Table 2 – Correlations between short timestep predictions and observed measurements

	Average Coefficient of Determination (R^2) with optimum lag for pollutant flux (TSS)	
Catchment (TSS)	Observed vs Predicted Loads (R^2)	Observed vs Predicted Concentration (R^2)
Mt. Waverley	0.72	0.46
Sheperds Rd	0.47	0.50
Narre Warren	0.64	0.70
Eley Rd	0.65	0.73
Richmond	0.85	0.58
Ruffeys	0.49	0.68

The temporal variation of event flow becomes the most important variable in short term prediction, and the difference in correlations between an average lag and an optimum lag demonstrates the sensitivity of predicting pollutants at such short timesteps. Table 2 shows however that if the lag can be predicted with accuracy; or if the exact timing of the peaks is not needed, that the model can achieve high correlations between observed and predicted points.

Table 3 – Relationship between W and Duration/Cumulative Volume

		W and Duration		**W andCumulative Volume**	
TSS	**No. Observations**	**Slope**	**p**	**Slope**	**p**
Mt. Waverley	509	-0.0011	0.01	0.0000	0.07
Sheperds Rd	251	0.0002	0.51	0.0000	0.44
Narre Warren	147	-0.0004	0.65	-0.0001	0.43
Eley Rd RB, Burwood East	191	-0.0005	0.16	0.0000	0.75

Madden Grove, Richmond	377	-0.0008	0.04	0.0000	0.04
Ruffeys Wetland, Doncaster	381	0.0006	0.00	0.0000	0.02
TP					
Mt. Waverley	299	-0.00031	0.43	0.010505	0.08
Sheperds Rd	134	0.00098	0.23	0.000018	0.23
Madden Grove, Richmond	185	-0.00386	0.01	-0.000007	0.33
TN					
Mt. Waverley	285	-0.000346	0.18	0.000004	0.10
Sheperds Rd	163	0.000013	0.45	0.000013	0.06
Madden Grove, Richmond	197	-0.001654	0.00	-0.000003	0.26

When all points, from all catchments, are considered together there is little indication of depletion from any of the catchments with either time elapsed or storm volume (Table 3). Although there are some catchments with a significant slope on the trendline, there is no clear preference for a declining or increasing *W*. The lack of a clear pattern suggests that there is no need to incorporate a depletion term in the model, consistent with a number of other studies (Duncan 1995; Chiew, Duncan et al. 1997; Smith 1997).

Conclusion

The proposed model can predict loads of total suspended solids, total phosphorus, and total nitrogen exported from an urban catchment with an explanatory power comparable to that for modelling event runoff volume from rainfall. The rainfall function is simple requiring only two fitted parameters – a power term describing the shape of the relationship, and a coefficient defining its magnitude. A physical explanation of the model centres on the energy required to maintain particulate contaminants in suspension, suggesting that this is the critical process.

The main advantage of this approach is in its simplicity and practicality, including the following:

- For the implementation of the model, only high resolution data on rainfall (that are usually widely available) are needed, while flow data is not required (usually very costly to obtain). The use of rainfall rather than runoff to estimate event loads avoids the direct spurious correlations between runoff and load, as noted by Huber (1992), allowing a legitimate predictive relationship to be used (Praire and Bird 1989).

- The model has only two calibration parameters that is a far less number than usually used in other approaches (e.g. very often 4 calibration parameters are needed). The clear advantage will be if we could demonstrate that the parameters are transferable from one site to another and therefore allow for modelling of non-monitored catchments.

Sensitivity analysis shows that the calibration parameters form unique solutions, although there is a relative insensitivity to small changes in the exponent in some cases, particularly on catchments with less representative data sets. Importantly the model appears to predict the intra event behaviour, but the analysis demonstrates that the timing of these predictions is critical, and those lags times vary considerably even within small catchments.

The simplicity and explanatory power of the model make it a prime candidate for widespread use on the mixed use urban catchments tested, rather like a rational method for loads, and a useful step towards formulation of a physically based, usable model.

Bibliography

Aalderink, R. H., E. H. S. Van Duin, et al. (1990). Some Characteristics of Run-off Quality from a Separated Sewer System in Lelystad, The Netherlands. Proceedings of the Fifth International Conference on Urban Storm Drainage, Osaka, Japan.

Ahyerre, M., G. Chebbo, et al. (1998). "Stormwater quality modelling, an ambitious objective?" Water Science Technology **37**(1): 205-213.

Akan, A. O. (1987). "Pollutant Washoff by Overland Flow." Journal of Environmental Engineering **113**(4): 811-823.

Alley, W. M., F. W. Ellis, et al. (1980). Toward a More Deterministic Urban Runoff-Quality Model. Proceedings of the International Symposium on Urban Storm Runoff, Lexington, Kentucky.

Alley, W. M. and P. E. Smith (1981). "Estimation of Accumulation Parameters for Urban Runoff Quality Modelling." Water Resources Research **17**(6): 1657-1664.

Allison, R. A., F. H. S. Chiew, et al. (1997). Stormwater Gross Pollutants. Melbourne, CRC Catchment Hydrology.

Arnberg-Nielson, K. and P. Harremoes (1996). "The importance of inherent uncertainties in state-of-the-art urban storm drainage modelling for ungauged small catchments." Journal of Hydrology **179**: 305-319.

Asplund, R., B. W. Mar, et al. (1982). "Total Suspended Solids in Highway Runoff in Washington State." Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers **108**(EE2): 391-404.

Baffaut, C. and J. W. Delleur (1990). "Calibration of SWMM Runoff Quality Model with Expert System." Journal of Water Resources Planning and Management **116**(2): 247-261.

Ball, J. E., R. Jenks, et al. (1998). "An assessment of the availability of pollutant constituents on road surfaces." The Science of the Total Environment **209**: 243-254.

Bannerman, R. T., D. W. Owens, et al. (1993). "Sources of Pollutants in Wisconsin Stormwater." Water Science and Technology **28**(3-5): 241-259.

Barfield, B. J., T. Y. Kao, et al. (1978). Sediment Control from Urban Construction Areas in the U.S. Urban Storm Drainage, Proceedings of the International Conference held at the University of Southampton, Pentech Press.

Beck, M. B. (1987). "Water quality modeling: A review of the analysis of uncertainty." Water Resources Research **23**(8): 1393-1442.

Bedient, P. B., J. L. Lambert, et al. (1980). "Stormwater Pollutant Load-Runoff Relationships." Journal of the Water Pollution Control Federation **52**(9): 2396-2404.

Berndtsson, R. and J. Niemczynowicz (1988). "Spatial and temporal scales in rainfall analysis - some aspects and future perspectives " Journal of Hydrology **100**: 293-313.

Bertrand-Krajewski, J. L. and J. P. Bardin (2002b). Uncertainties and representativity of measurements in stormwater storage tanks. 9th ICUD, Portland, Portland, Oregon, USA.

Bertrand-Krajewski, J. L., S. Barraud, et al. (2002). "Uncertainties, performance indicators and decision aid applied to stormwater facilities." Urban Water **4**: 163-179.

Bertrand-Krajewski, J. L., G. Chebbo, et al. (1998). "Distribution of pollutant mass vs volume in stormwater discharges and the first flush phenomenon." Water Research **32**(8): 2341-2356.

Bertrand-Krajewski, J. L. and O. Scrivener (1993). "Sewer sediment production and transport modelling: A literature review." Journal of Hydraulic Research **31**(4): 435-460.

Bertrand-Krajewski, J. L., O. Scrivener, et al. (1993). "Sewer sediment production and transport modelling: A literature review." Journal of Hydraulic Research **31**(4): 435-460.

Boyd, M. J., M. C. Bufill, et al. (1994). "Pervious and impervious runoff in urban catchments." Hydrological Sciences Journal **38**(6): 463-478.

Bradford, W. L. (1977). "Urban Stormwater Pollutant Loadings: A Statistical Summary Through 1972." Journal of the Water Pollution Control Federation **49**(4): 613-622.

Brezonik, L. and T. H. Stadelmann (2002). "Analysis and predictive models of stormwater runoff volumes, loads, and pollutant concentrations from watersheds in the Twin Cities metropolition area, Minnesota, USA." Water Research **36**: 1743-1757.

Brodie, I. (2007). "Prediction of stormwater particle loads from impervious urban surfaces based on a rainfall detachment index." Water Science and Technology **55**(4): 49-55.

Butler, D., S. Thedchanamoorthy, et al. (1992). "Aspects of Surface Sediment Characteristics on an Urban Catchment in London." Water Science and Technology **25**(8): 13-19.

Butts, M. B., Payne, J.T., Kristensen, M., Madsen, H., (2004). "An evaluation of the impact of model structure on hydrological modelling uncertainty for streamflow simulation." Journal of Hydrology **298**: 242-266.

Calabro, P. S. (2001). "Cosmoss:conceptual simplified model for sewer system simulation A new model for urban runoff quality." Urban Water **3**: 33-42.

Castaldi, F. J. (1983). Final Report of the Nationwide Urban Runoff Program in Austin, Texas. Austin, USEPA.

Characklis, W. G., P. B. Bedient, et al. (1978). "A Study of Stormwater Runoff Quality." Progress in Water Technology **10**(5/6): 673-699.

Charbeneau, R. J. and M. E. Barrett (1998). "Evaluation of Methods for Estimating Stormwater Pollutant Loads." Water Environment Research **70**(7): 1295-1302.

Chaubey, I., C. T. Haan, et al. (1999). "Uncertainty in the model parameters due to spatial variability of rainfall." Journal of Hydrology **220**: 48-61.

Chiew, F. H. S., H. P. Duncan, et al. (1997). Modelling Pollutant Buildup and Washoff: Keep It Simple. Wai-Whenua 1997, 24th Hydrology and Water Resources Symposium, Auckland, New Zealand, New Zealand Hydrological Society.

Chiew, F. H. S. and T. A. McMahon (1998). Estimation of Stormwater Runoff and Diffuse Pollution Loads. HydraStorm 98, 3rd International Symposium on Stormwater Management, Adelaide, South Australia, IEAust.

Choi, K. and J. E. Ball (2002). "Parameter estimation for urban runoff modelling." Urban Water(4): 31-41.

Chui, T. W., B. W. Mar, et al. (1982). "Pollutant Loading Model for Highway Runoff." Journal of the Environmental Engineering Division, Proceedings of the American Society of Civil Engineers **108**(EE6): 1193-1210.

City Design (2003). Stormwater Monitoring Program 2002-2003. Brisbane, City Design/Brisbane City Council.

Clark, S., Pitt, R., Burian, S., Field, S., Fan, E., Heaney, J., Wright, L. (2006). Annotated bibliography of wet weather flow literature from1996 through 2004. Middleton, PA.

Coleman, T. J. (1993). A Comparison of the Modelling of Suspended Solids Using SWMM3 Quality Prediction Algorithms with a Model Based on Sediment Transport Theory. Proceedings of the Sixth International Conference on Urban Storm Drainage, Niagara Falls, Ontario, Canada.

Cordery, I. (1977). "Quality Characteristics of Urban Stormwater in Sydney, Australia." Water Resources Research **13**(1): 197-201.

Cowen, W. F. and G. F. Lee (1976). "Phosphorus Availability in Particulate Materials Transported by Urban Runoff." Journal of the Water Pollution Control Federation **48**(3): 580-591.

CRC for Freshwater Ecology (2004). Urban stormwater and the ecology of streams. Canberra, Cooperative Research Centre for Freshwater Ecology.

CRCCH (2003). MUSIC users guide: Version 2, Co-operative Research Centre for Catchment Hydrology.

Cristina, C. M. and J. J. Sansalone (2003). ""First Flush", power law and particle seperation diagrams for urban storwater suspended particulates." Journal of Environmental Engineering **April 2003**: 298-307.

Cristina, C. M. and J. J. Sansalone (2003). "‘‘First Flush,’’ Power Law and Particle Separation Diagrams for Urban Storm-Water Suspended Particulates." Journal of Environmental Engineering **April 2003**: 298-307.

CSIRO (1996). Port Phillip Bay Environmental Study, CSIRO.

Deletic, A., C. Maksimovic, et al. (1997). "Modelling of Storm Washoff of Suspended Solids from Impervious Surfaces." Journal of Hydraulic Research **35**(1): 99-118.

Deletic, A. and D. W. Orr (2003). Water quality of runoff from paved surfaces. IAHR Congress, Thessaloniki, Greece.

Deletic, A. B. (1998). "Evaluation of Water Quality Factors in Storm Runoff from Paved Areas." Journal of Environmental Engineering **124**(9): 869-879.

Deletic, A. B. (1998). "The first flush load of urban surface runoff." Water Research **32**(8): 2462-2470.

Desbordes, M. and E. Servat (1984). Solids in Urban Runoff - Statistical Analysis of French Experimental Data. Proceedings of the Third International Conference on Urban Storm Drainage, Goteborg, Sweden.

Desbordes, M. and E. Servat (1984). Solids in Urban Runoff: Statistical Analysis of French Experimental Data. 3rd International Conference on USD. 1984 Goteborg, Goteborg.

Desbordes, M. and E. Servat (1987). Transport Model over Urban Catchment Surfaces Regarding TSS, BOD5, and COD. Urban Storm Water Quality, Planning and Management, Proceedings of the Fourth International Conference on Urban Storm Drainage, Lausanne, Switzerland.

Deutsch, J. C. and M. Desbordes (1981). Study of Runoff Pollution for Urban Planning. Urban Stormwater Quality, Management, and Planning, Proceedings of the Second International Conference on Urban Storm Drainage, Urbana, Illinois, Water Resources Publications, Colorado.

Deutsh, J. C. and J. C. Hemain (1984). Main results of the French National Programme of Urban Runoff Quality Measurement. 3rd International Conference on USD. 1984 Goteborg, Goteborg.

Dotto, C., Deletic, A., Fletcher, T.D., (2009). "Analysis of parameter uncertainty of a flow and quality stormwater model." Water Science and Technology **60**(3): 717-725.

Driscoll, E. D. (1986). Lognormality of Point and Non-Point Source Pollutant Concentrations. Urban Runoff Quality - Impact and Quality Enhancement Technology, Proceedings of an Engineering Foundation Conference, New Hampshire, ASCE.

Driver, N. E. (1990). Summary of Nationwide Analyses of Storm-Runoff Quality and Quantity in Urban Watersheds. Proceedings of the Fifth International Conference on Urban Storm Drainage, Osaka, Japan.

Driver, N. E. and B. M. Troutman (1989). "Regression Models for Estimating Urban Storm-Runoff Quality and Quantity in the United States." Journal of Hydrology **109**: 221-236.

Duchon, C. E. and G. R. Essenberg (2001). "Comparative rainfall observations from pit and aboveground rain gauges with and without wind shields." Water Resources Research **37**(12): 3253-3263.

Duda, A. M., D. R. Lenat, et al. (1982). "Water Quality in Urban Streams - What We Can Expect." Journal of the Water Pollution Control Federation **54**(7): 1139-1147.

Duncan, H. P. (1995). A Review of Urban Storm Water Quality Processes. Melbourne, Australia, Cooperative Research Centre for Catchment Hydrology.

Duncan, H. P. (1995a). A Bibliography of Urban Storm Water Quality. Melbourne, Australia, Cooperative Research Centre for Catchment Hydrology.

Duncan, H. P. (1999). Urban Stormwater Quality: A Statistical Overview. Melbourne, Australia, Cooperative Research Centre for Catchment Hydrology.

Ellis, J. B. (1986). Pollutional aspects of urban runoff. Urban Runoff Pollution. H. C. Torno, J. Marsalek and M. Desbordes. Berlin, Springer-Verlag. **G10:** 1-38.

Engineers Australia (1997). Australian Rainfall and Runoff Crows Nest Engineers Australia.

Fergusson, J. E. and D. E. Ryan (1984). "The Elemental Composition of Street Dust from Large and Small Urban Areas Related to City Type, Source and Particle Size." The Science of the Total Environment **34**: 101-116.

Fowler, J., L. Cohen, et al. (1998). Practical Statistics for Field Biology. Chichester, England, John Wiley and Sons.

Francey, M., A. Deletic, et al. (2004). Monitoring of Stormwater Runoff in Melbourne. 19th European Junior Scientists Workshop - Process Data and Integrated Urban Water Modelling, Lyon, France.

Freund, A. P. and C. D. Johnson (1980). Comparison and Relationships of Stormwater Quality and Basin Characteristics: Madison, Wisconsin. Proceedings of the International Symposium on Urban Storm Runoff, Lexington, Kentucky.

Fuchs, S., H. Brombach, et al. (2004). New database on urban runoff pollution. Novatech, Lyon, France.

Gaume, E., J. Villeneuve, et al. (1998). "Uncertainty assessment and analysis of the calibrated parameter values of an urban storm water quality model." Journal of Hydrology **210**: 38-50.

Geiger, W. (1987). Flushing effects in combined sewer systems. 4th International Conference on urban Drainage, Lausanne, Switzerland.

Geiger, W. F. (1984). Goodness and Limits in Urban Runoff Modelling. Proceedings of the Third International Conference on Urban Storm Drainage, Goteborg, Sweden.

GeoSyntec Consultants and ASCE (2002). Urban Stormwater BMP Performance Monitoring. Washington DC, US EPA.

GHD and EPA Vic (1981). Characterisation of Pollution in Urban Stormwater Runoff. Canberra, Australian Water Resources Council.

Ginzburg, L. R. and C. X. J. Jensen (2004). "Rules of thumb for judging ecological theories." Trends in Ecology and Evolution **19**(3): 121-126.

Grayson, R., I. D. Moore, et al. (1992). "Physically based hydrologic modelling 2. Is the concept realistic?" Water Resources Research **26**(10): 2659-2666.

Grottker, M. (1987). "Runoff Quality from a Street with Medium Traffic Loading." The Science of the Total Environment **59**: 457-466.

Gutteridge Haskins and Davey and Environment Protection Authority of Victoria (1981). Characterisation of Pollution in Urban Stormwater Runoff. Canberra, Australian Water Resources Council.

Hammer, T. R. (1972). "Stream Enlargement due to Urbanisation." Water Resources Research **8**(6): 1530-1540.

Harmel, R. D., Cooper, R.J., Slade, R.M., Haney, R.L., Arnold, J.G., (2006). "Cumulative uncertainty in measured streamflow and water quality data for small watersheds." Transactions of the American Society of Agricultural and Biological Engineers **43**(3): 689-701.

Harmel, R. D., King, K. W., Wolfe,J. E., Torbert, H. A (2002). "Minimum flow considerations for automated storm sampling on small watersheds." Texas Journal of science **54**(2): 177-188.

Harrison, R. M. (2002). "Key Pollutants - Airborne Particles." Clean Air and Environmental Protection **32**(4): 38-41.

Hartigan, J. P., A. M. Lumb, et al. (1978). Calibration of Urban Nonpoint Pollution Loading Models. Verification of Mathematical and Physical Models in Hydraulic Engineering, Proceedings of the 26th Annual Hydraulics Division Specialty Conference, University of Maryland, Hydraulics Division, ASCE, New York.

Haster, T. W. and W. P. James (1994). "Predicting Sediment Yield in Storm-Water Runoff from Urban Areas." Journal of Water Resources Planning and Management **120**(5): 630-650.

Hatt, B. E., T. D. Fletcher, et al. (2004). "The Influence of Urban Density and Drainage Infrastructure on the Concentrations and Loads of Pollutants in Small Streams." Environmental Management **34**(1): 112-124.

House, M. A., J. B. Ellis, et al. (1993). "Urban Drainage - Impacts on Receiving Water Quality." Water Science and Technology **27**(12): 117-158.

Huber, W. C. (1986). Modelling Urban Runoff Quality: State of the Art. Urban Runoff Quality - Impact and Quality Enhancement Technology, Proceedings of an Engineering Foundation Conference, New Hampshire, ASCE, New York.

Huber, W. C. (1992). Prediction of Urban Nonpoint Source Water Quality: Methods and Models. International Symposium on Urban Stormwater Management, Sydney.

Huber, W. C. and R. E. Dickinson (1988). Storm Water Management Model, Version 4: Users' Manual. Gainesville, University of Florida.

Huber, W. C. and J. P. Heaney (1980). Operational Models for Stormwater Quality Management. Environmental Impact of Nonpoint Source Pollution. M. J. Overcash and J. M. Davidson, Ann Arbor Science Publishers Inc**:** 397-444.

Hudson, N. (1971). Soil Conservation. London, B T Batsford Ltd.

Hvitved-Jacobsen, T., Y. A. Yousef, et al. (1984). "Fate of Phosphorus and Nitrogen in Ponds Receiving Highway Runoff." The Science of the Total Environment **33**: 259-270.

Irish, L. B., M. E. Barrett, et al. (1998). "Use of Regression Models for Analyzing Highway Stormwater Loads." Journal of Environmental Engineering(October 1998): 987-993.

Jakeman, A. J. and G. M. Hornberger (1993). "How Much Complexity is Warranted in a Rainfall-Runoff Model." Water Resources Research **29**(8): 2637-2649.

Jakeman, A. J., J. A. Taylor, et al. (1992). "Efficient Tools for Analysising the Influence of Sources and Meteorology on Urban Ambient COncentration Trends Illustrated for Canberra, Australlia." Ecological Modelling **64**: 125-157.

James, W. and B. Shivalingaiah (1986). Continuous Mass Balance of Pollutant Build-up Processes. Urban Runoff Pollution. H. C. Torno, J. Marsalek and M. Desbordes. Berlin, Springer-Verlag. **G10:** 243-271.

Jewell, T. K. and D. D. Adrian (1982). "Statistical Analysis to Derive Improved Stormwater Quality Models." Journal of the Water Pollution Control Federation **54**(5): 489-499.

Jones, J. E. (1986). Urban Runoff Impacts on Receiving Waters. Urban Runoff Quality - Impact and Quality Enhancement Technology, Proceedings of an Engineering Foundation Conference, New Hampshire, ASCE.

Kanso, A., M. C. Gromaire, et al. (2003). "Bayesian approach for the calibration of models: Application to an urban stormwater polltion model." Water Science and Technology **47**(4): 77-84.

Kanso, A., B. Tassin, et al. (2004). Managing uncertainties in urban runoff quality models: A benching methodolgy. Novatech, Lyon, France.

Kanso, A., B. Tassin, et al. (2004). Managing uncertainties in urban runoff quality models: A benchmarking methodolgy. Novatech. Lyon, France.

Kenney, B. C. (1982). "Beware of spurious self-correlations." Water Resources Research **18**(4): 1041-1048.

Kuczera, G., Kavetski, D., Franks, S., Thyer, M., (2006). "Towards a bayesian total error analysis of conceptual rainfall-runoff models: characterising model error using storm-dependent parameters." Journal of Hydrology **331**(1-2): 161-177.

Kuo, J. T., Y. L. Yan, et al. (1993). "A Simplified Computer Model for Nonpoint Source Pollution in a Small Urban Area." Water Science and Technology **28**(3-5): 701-706.

LaBarbera, P., L. G. Lanza, et al. (2002). "Tipping bucket mechanical errors and their influence on rainfall statistics and extremes." Water Science and Technology **45**(2): 1-9.

Lee, D. S., J. A. Garland, et al. (1994). "Atmospheric Concentrations of Trace Elements in Urban Areas of the Unitied Kingdom." Atmos Environ **28**(16): 2691-2713.

Lee, J. H., K. W. Bang, et al. (2002). "First flush analysis of urban storm runoff." The Science of the Total Environment(293): 163-175.

Leecaster, M. K., K. Schiff, et al. (2002). "Assessment of efficient sampling designs for urban stormwater monitoring." Water Research **36**: 1556-1564.

Legates, D. R. and G. J. McCabe (1999). "Evaluating the use of "goodness-of-fit" measures in hydrologic and hydroclimatic model validation." Water Resources Research **35**(1): 233-241.

Lloyd, S. D. (2004). Exploring the opportunities and barriers to sustainable stormwater management practices in residential catchments. Department of Engineering. Melbourne, Monash.

Maksimovic, C. and Radojkovic (1986). Urban Drainage Catchments: Selected worldwide rainfall-runoff data from experimental catchments., Pergamon Press, Belgrade.

Malmquist, P. (1978). "Atmospheric Fallout and Street Cleaning - Effects on Urban Storm Water and Snow." Progress in Water Technology **10**(5/6): 495-505.

Mance, G. and M. M. I. Harman (1978). The Quality of Urban Storm-Water Runoff. Urban Storm Drainage, Proceedings of the International Conference held at the University of Southampton, Southampton, Pentech Press, London.

MapInfo Corporation (2002). MapInfo Professional Version 7.0.

Marsalek, J. (1981). "Calibration of the tipping-bucket raingauge." Journal of Hydrology **53**: 343-354.

Mason, Y., A. A. Ammann, et al. (1999). "Behaviour of heavy metals, nutrients, and major components during roof runoff infiltration." Environmental Science Technology **33**: 1588-1597.

McCarthy, D. T., Deletic, A., Mitchell, V.G., Fletcher, T.D., Diaper, C., (2008). "Uncertaities in stormwater E. coli levels." Water Research **42**(6-7): 1812-1824.

McPherson, T. N., S. J. Burian, et al. (2002). "Comparison of the pollutant loads in dry and wet weather runoff in a southern California urban watershed." Water Science and Technology **45**(9): 225-261.

McPherson, T. N., Burian, S. J., Turin, H.J., Strenstron, M.K., Suffet, I.H., (2002). "Comparison of the pollutant loads in dry and wet weather runoff in a southern California urban watershed
" Water Science and Technology **45**(9): 255-261.

Melways Publishing (2002). Melways Greater Melbourne Street directory. Melbourne, Melways Publishing Pty. Ltd.

Meyer, L. D., Wischmeier, W.H., (1969). "Mathematical simulation of the process of soil erosion by water." Trans. Amer. Soc. Agric. Eng **12**: 756-762.

Miller, R. A. and H. C. Mattraw (1982). "Storm Water Runoff Quality from Three Land-Use Areas in South Florida." Water Resources Bulletin **18**(3): 513-519.

Molini, A., L. G. Lanza, et al. (2005). "The impact of tipping-bucket raingauge measurement errors on design rainfall for urban-scale applications." Hydrological Processes **19**: 1073-1088.

Moreira, L. F. F., A. M. Righetto, et al. (2006). Uncertainty analysis associated with Rainfall Spatial Distribution in an Experimental Semiarid Watershed, Northeastern Brazil. 3rd Biennial meeting of the International Environmental Modelling and Software Society. Vermont, USA.

Muthukaruppan, M., F. H. S. Chiew, et al. (2002). Size Distribution and Partitioning of Urban Pollutants. 9th International Conference on Urban Drainage, Portland, Oregan, USA, ASCE.

Nakamura, E. (1984). Factors Affecting Stormwater Quality Decay Coefficient. Proceedings of the Third International Conference on Urban Storm Drainage, Goteborg, Sweden.

Nash, J. E. and J. V. Sutcliffe (1970). "River forecasting through conceptual models, 1. A discussion of principles." Journal of Hydrology **10**: 282-290.

Nelen, F., A. Mooijman, et al. (1992). "The importance of rainfall distribution in urban rainfall operation." Nordic Hydrology **23**: 121-136.

North Shore City Council (2002). Project CARE: 2001 Monitoring. Auckland, North Shore City Council.

Novotny, V., H.-M. Sung, et al. (1985). "Estimating Nonpoint Pollution from Small Urban Watersheds." Journal of the Water Pollution Control Federation **57**(4): 339-348.

Odnevall Wallinder, W. H. I., Leygraf, C., (2000). "A laboratory study of copper and zinc runoff during first flush and steady state conditions." Corrosion Science **43**(1): 127-146.

Ogden, F. L. and P. Julien, Y. (1993). "Runoff Sensitivity to Temporal and Spacial Rainfall Variability at Runoff Plane and Small Basin Scales." Water Resources Research **29**(8): 2589-2597.

Orr, D. W. (2002). Pollution Release from Urban Asphalt Surfaces, Aberdeen University.

Orr, D. W. and A. B. Deletic (2003). Pollution Build up on road surfaces. Aberdeen.

Pitt, R. (1979). Demonstration of Nonpoint Pollution Abatement Through Improved Street Cleaning Practices. Cincinnati, United States Environmental Protection Agency.

Pitt, R. and A. Maestre (2005). Stormwater quality as described in the National Stormwater Quality Database (NSQD). 10th ICUD, Copenhagen, Denmark.

Prairie, Y. T. and D. F. Bird (1989). "Some misconceptions about the spurious correlation problem in the ecological literature." Oecologia **81**: 285-288.

Price, R. K. and G. Mance (1978). A Suspended Solids Model for Storm Water Runoff. Urban Storm Drainage, Proceedings of the International Conference held at the University of Southampton, Pentech Press.

Randall, C. W., K. Ellis, et al. (1982). Urban Runoff Pollutant Removal by Sedimentation. Stormwater Detention Facilities - Planning, Design, Operation, and Maintenance, Henniker, New Hampshire, ASCE.

Randall, C. W., T. J. Grizzard, et al. (1981). Comparison of Pollutant Mass Loads in Precipitation and Runoff in Urban Areas. Urban Stormwater Quality, Management, and Planning, Proceedings of the Second International Conference on Urban Storm Drainage, Urbana, Illinois, Water Resources Publications, Colorado.

Reinertsen, T. R. (1981). Quality of Stormwater Runoff from Streets. Urban Stormwater Quality, Management, and Planning, Proceedings of the Second International Conference on Urban Storm Drainage, Urbana, Illinois, Water Resources Publications, Colorado.

RossRakesh, S., C. Gippel, et al. (1999). Blackburn Lake discharge and water quality monitoring program:Data summary and interpretation. Melbourne, CRC for Catchment Hydrology/Freshwater Ecology.

Salles, C., J. Poesen, et al. (2000). "Statistical and Physical Analysis of Soil Detachment by Raindrop Impact: Rain Erosivity Indices and Threshold Energy." Water Resources Research **36**(9): 2721-2729.

Sartor, J. D. and G. B. Boyd (1972). Water Pollution Aspects of Street Surface Contaminants. Washington, D.C., United States Environmental Protection Agency.

Sartor, J. D., G. B. Boyd, et al. (1974). "Water Pollution Aspects of Street Surface Contaminants." Journal of the Water Pollution Control Federation **46**(3): 458-467.

Shaheen, D. G. (1975). Contributions of Urban Roadway Usage to Water Pollution. Washington, D.C., United States Environmental Protection Agency.

Shapiro, S. S. and M. B. Wilk (1965). "An analysis of variance test for normality." Biometrika **52**(3): 591-9.

Sharpin, M. G. (1993). Hydrological Variability, Urbanisation and Water Quality in Australia. Proceedings of the Sixth International Conference on Urban Storm Drainage, Niagara Falls, Ontario, Canada.

Sharpin, M. G. (1995). Stormwater Quality Characteristics from Urban and Non-Urban Catchments in South-Eastern Australia. AWWA 16th Federal Convention, Darling Harbour, Sydney, Australia, Australian Water and Wastewater Association.

Shivalingaiah, B. and W. James (1984a). Applied Algorithms for Pollutant Buildup, Washoff and Transport in Stormwater. Water Quality Evolution Within the Hydrological Cycle of Watersheds, Proceedings of the Canadian Hydrology Symposium No 15, Quebec, Canada.

Shivalingaiah, B. and W. James (1984b). Algorithms for Buildup, Washoff, and Routing Pollutants in Urban Runoff. Proceedings of the Third International Conference on Urban Storm Drainage, Goteborg, Sweden.

Shuster, W. D., J. Bonta, et al. (2005). "Impacts of impervious surface on watershed hydrology: A review." Urban Water Journal **2**(4): 263-275.

Smith, W. (1997). Estimation of pollution accumulation for stormwater washoff modelling. Civil Engineering. Melbourne, Monash.

Smullen, J. T., A. L. Shallcross, et al. (1999). "Updating the U.S. nationwide urban runoff quality database." Water Science Technology **39**(12): 9-16.

Spangberg, A. and J. Niemczynowicz (1993). Measurements of Pollution Wash-off from an Asphalt Surface. Proceedings of the Sixth International Conference on Urban Storm Drainage, Niagara Falls, Ontario, Canada.

Sprent, J. I., Parsons, R. (2000). "Nitrogen fixation in legume and non-legume trees
" Field Crops Research **65**(2-2).

Sriananthakumar, K. and G. P. Codner (1992). Factors Affecting Pollutant Washoff Decay Coefficient in Urban Catchments. International Symposium on Urban Stormwater Management, Sydney, Australia.

Sriananthakumar, K. and G. P. Codner (1992a). Data Evaluation Before Urban Runoff Quality Modelling. International Symposium on Urban Stormwater Management, Sydney.

Sriananthakumar, S. K. and G. P. Codner (1993). A Process Based Model for Pollution Washoff by Overland Flow. Hydrology and Water Resources Symposium, Newcastle, IEAust.

Strangeways, I. (2004). "Improving precipitation measurement." International Journal of Climatology **24**: 1443-1460.

Talylor, B. N., Kuyatt, C.E., (1994). Guidelines for evaluating and expressing the uncertainty of NIST measurement results. Technical Note 1296, National Institute of Standards and Technology, USA.

Tasker, G. D. and N. E. Driver (1988). "Nationwide Regression Models for Predicting Urban Runoff Water Quality at Unmetered Sites." Water Resources Bulletin **24**(5): 1091-1101.

Taylor, G. D., T. D. Fletcher, et al. (2005). "Nitrogen composition in urban runoff—implications for stormwater management." Water Research **39**: 1982-1989.

Troutman, B. M. (1983). "Runoff prediction erros and bias in parameter estimation nduced by spatial variability of rainfall on storm runoff." Water Resources Research **19**(3): 791-810.

Tsihrintzis, V. A. and R. Hamid (1998). "Runoff quality prediction from small urban catchments using SWMM." Hydrological Processes **12**: 311-329.

van Dijk, A. I. J. M., L. A. Bruijnzeel, et al. (2002). "Rainfall intensity - kinetic energy relationships:a critical literature appraisal." Journal of Hydrology **261**(2002): 1-23.

Var, F., Y. Narita, et al. (2000). "The Concentration, Trend and Seasonal Variation of Metals in the Atmosphere in 16 Japenese Cities shown by the Results of National Air Surveillance Network (NASN) from 1974 to 1996." Atmos Environ **34**: 2755-2770.

Vaze, J. and F. H. S. Chiew (2002). "Experimental study of pollutant accumulation on an urban road surface." Urban Water **4**(2002): 379-389.

Vaze, J. and F. H. S. Chiew (2003). "Comparative evaluation of urban storm water quality models." Water Resources Research **39**(10): 1280.

Vaze, J. and F. H. S. Chiew (2003b). "Study of pollutant washoff from small impervious experimental plots." Water Resources Research **39**(6): 1160.

Vorreiter, L. and C. Hickey (1994). Incidence of the First Flush Phenomenon in Catchments of the Sydney Region. Water Down Under 94, Adelaide, South Australia.

Walden, W. J. (1999). Non-point source pollutant export estimation from urban catchments. Faculty of Engineering. Brisbane, University of Queensland.

Walsh, C. J., T. Fletcher, et al. (2004). Urban stormwater and the ecology of streams. Canberra, Cooperative Research Centre for Freshwater Ecology.

Walsh, C. J., A. K. Sharpe, et al. (2001). "Effects of Urbanization on Streams of the Melbourne Region, Victoria, Australia. I. Benthic Macroinvertebrate Communities." Freshwater Biology(46).

Wang, X., S. Tong, et al. (2006). "Air Pollution Trends in Brisbane, Australia, between 1980 and 2003." Clean Air and Environmental Quality **40**(1): 34-39.

Weibel, S. R., R. J. Anderson, et al. (1964). "Urban Land Runoff as a Factor in Stream Pollution." Journal of the Water Pollution Control Federation **36**(7): 915-924.

Wong, T. H. F., P. F. Breen, et al. (2000). Water Sensitive Road Design options for improving stormwater quality of road runoff, Cooperative research centre for catchment hydrology.

Xanthopoulos, C. and H. H. Hahn (1993). Anthropogenic Pollutants Wash-off from Street Surfaces. Proceedings of the Sixth International Conference on Urban Storm Drainage, Niagara Falls, Ontario, Canada.

Yaziz, M. I., H. Gunting, et al. (1989). "Variations in Rainwater Quality from Roof Catchments." Water Research **23**(6): 761-765.

Young, R. A. and J. L. Wiersma (1973). "The role of rainfall impact in soil detachment and transport." Water Resources Research **9**(6): 1629-1636.

Yuan, Y., K. Hall, et al. (2001). "A preliminary model for predicting heavy metal contaminent loading from an urban catchment." The Science of the Total Environment **266**: 299-307.

Zoppou, C. (2001). "Review of urban storm water models." Environmental Modelling & Software **16**: 195-231.