Dispersed Urban-Stormwater Control Improved Stream Water Quality in a Catchment-Scale Experiment

Christopher J Walsh[1], Moss J Imberger[1], Matthew J Burns[1], Tim D Fletcher[1], and Darren G Bos[2]

[1]University of Melbourne
[2]The University of Melbourne

November 22, 2022

Abstract

Traditional urban drainage degrades receiving waters. Alternative approaches have potential to protect downstream waters, but widespread adoption requires robust demonstration of their feasibility and effectiveness. We conducted a catchment-scale experiment over 19 years to assess the effect of dispersed stormwater control measures (SCMs), measured as a reduction in effective imperviousness (EI) on stream water quality in 6 sites on 2 streams. We compared changes in those sites over 7 years as EI decreased, to changes in the 12 preceding years, and in 3 reference and 2 control streams. SCMs reduced phosphorus concentrations and summer temperature to reference levels in dry weather where EI was sufficiently reduced, but effects were smaller with increased antecedent rain. SCMs also reduced nitrogen concentrations which were influenced by septic tank seepage in all sites. SCMs had no effect on suspended solids concentrations, which were lower in urban than in reference streams. SCMs increased electrical conductivity: along with reduced temperature this is evidence of increased contribution of groundwater to baseflows. This experiment strengthens inference that urban stormwater drainage increases contaminant concentrations in streams, and demonstrates that such impacts are reversible and likely preventable. Variation in degree of water quality improvement among experimental sites suggests that achieving reference water quality would require SCMs with large retention capacity intercepting runoff from nearly all impervious surfaces, thus requiring more downslope space and water demand. EI is a useful metric for predicting stream water quality responses to SCMs, allowing better catchment prioritization and SCM design standards for stream protection.

Hosted file

`walsh_etal_suppl_info.docx` available at https://authorea.com/users/56953/articles/598989-dispersed-urban-stormwater-control-improved-stream-water-quality-in-a-catchment-scale-experiment

Hosted file

`essoar.10510253.1.docx` available at https://authorea.com/users/56953/articles/598989-dispersed-urban-stormwater-control-improved-stream-water-quality-in-a-catchment-scale-experiment

Christopher J. Walsh[1], Moss J. Imberger[1], Matthew J. Burns[1], Darren G. Bos[1], and Tim D. Fletcher[1]

[1] School of Ecosystem and Forest Sciences, The University of Melbourne, 500 Yarra Boulevard, Burnley, 3121 Victoria, Australia.

Corresponding author: Christopher J. Walsh (cwalsh@unimelb.edu.au).

†Additional author notes should be indicated with symbols (current addresses, for example).

Key Points:

- A long-term catchment-scale experiment strengthens inference that urban stormwater drainage increases stream contaminant concentrations.
- Dispersed stormwater control measures can reverse stormwater-induced degradation of stream water quality.
- Achieving reference stream water quality requires retention, treatment and loss of runoff from nearly all catchment impervious surfaces.

Author Contributions:

Conceptualization: Christopher J. Walsh and Tim D. Fletcher.

Data curation: Christopher J. Walsh, Moss J. Imberger, and Darren G. Bos.

Formal analysis: Christopher J. Walsh, Moss J. Imberger, Matthew J. Burns, and Tim D. Fletcher.

Funding acquisition: Christopher J. Walsh, Tim D. Fletcher, and Darren G. Bos.

Investigation: Christopher J. Walsh, Moss J. Imberger, Matthew J. Burns, Tim D. Fletcher, and Darren G. Bos.

Methodology: Christopher J. Walsh, Moss J. Imberger, Matthew J. Burns, and Tim D. Fletcher.

Project administration: Christopher J. Walsh, Tim D. Fletcher, and Darren G. Bos.

Software: Christopher J. Walsh, Moss J. Imberger, and Matthew J. Burns.

Validation: Christopher J. Walsh, Moss J. Imberger, Tim D. Fletcher, and Matthew J. Burns.

Visualization: Christopher J. Walsh.

Writing - original draft: Christopher J. Walsh and Moss J. Imberger.

Writing - review & editing: Christopher J. Walsh, Moss J. Imberger, Matthew J. Burns, Tim D. Fletcher, and Darren G. Bos.

Abstract

Traditional urban drainage degrades receiving waters. Alternative approaches have potential to protect downstream waters, but widespread adoption requires robust demonstration of their feasibility and effectiveness. We conducted a catchment-scale experiment over 19 years to assess the effect of dispersed stormwater control measures (SCMs), measured as a reduction in effective imperviousness (EI) on stream water quality in 6 sites on 2 streams. We compared changes in those sites over 7 years as EI decreased, to changes in the 12 preceding years, and in 3 reference and 2 control streams. SCMs reduced phosphorus concentrations and summer temperature to reference levels in dry weather where EI was sufficiently reduced, but effects were smaller with increased antecedent rain. SCMs also reduced nitrogen concentrations which were influenced by septic tank seepage in all sites. SCMs had no effect on suspended solids concentrations, which were lower in urban than in reference streams. SCMs increased electrical conductivity: along with reduced temperature this is evidence of increased contribution of groundwater to baseflows. This experiment strengthens inference that urban stormwater drainage increases contaminant concentrations in streams, and demonstrates that such impacts are reversible and likely preventable. Variation in degree of water quality improvement among experimental sites suggests that achieving reference water quality would require SCMs with large retention capacity intercepting runoff from nearly all impervious surfaces, thus requiring more downslope space and water demand. EI is a useful metric for predicting stream water quality responses to SCMs, allowing better catchment prioritization and SCM design standards for stream protection.

Plain Language Summary

The way we drain our cities and towns pollutes and erodes our streams and rivers. Water running off, and heated by, roofs and roads carries damaging particles and chemicals. The stormwater drains and pipes that we build transport the polluted runoff, quickly and untreated, to downstream waters. We installed 100s of rain-gardens that allow water soak into surrounding soils and be taken up by plants, and rainwater tanks for harvesting, in two suburban catchments to test if we could restore the water quality of creeks downstream. We compared before and after conditions in creeks downstream of our treatments, with conditions in other degraded urban streams, and in undegraded forested streams. Filtering and harvesting stormwater reduced summer temperatures and reduced concentrations of phosphorus and nitrogen, critical contaminants for healthy streams. The reductions were greatest in dry weather, and after small amounts of rain. To achieve water quality similar to forested streams, we need rain-gardens and harvesting systems that catch runoff from nearly every roof and road upstream. To achieve that, we need to put aside space near pipe outlets to streams for final treatment systems, and we need to find ways to use the excess water generated by roofs and roads.

1 Introduction

As the global human population has grown, associated demands on fresh water, and alteration to the world's lands, have resulted in growing water scarcity and declining water quality that threaten human societies and the freshwater ecosystems on which they depend (Vörösmarty et al., 2010). Cities and towns are the foci of such impacts, driving land use activities and degrading ecosystems well beyond their boundaries, in order to feed and water their populations, and to conduct their activities and build infrastructure (Grimm et al., 2008). Within their boundaries, conventional approaches to urban water management degrade rivers and streams. Streams of urban catchments are typified by flashy flow regimes, incised and deepened channels and compromised water quality leading to shifts in their ecological structure and function (Walsh, Roy, et al., 2005). Regionally, the impacts of urban land use on river water quality can be disproportionately large (e.g. Piffer et al., 2021). However, alternative land and water management practices have the potential to protect downstream waters and provide many co-benefits to the world's cities, including greater water supply security (Walsh et al., 2016). Their adoption for stream protection and restoration depends on robust demonstration of their feasibility and effectiveness.

Urban lands alter the water balance of catchments by replacing vegetated land with impervious surfaces, such as roofs and roads, which reduce rain-water infiltration and are usually efficiently drained to permit rapid routing of runoff to downstream waters (Walsh et al., 2012). Depending on the nature of local urban water infrastructure, these urban stormwater drainage effects may be exacerbated by wastewater discharges, by extraction of water from streams or aquifers that flow to them, or by channel engineering to mitigate flooding. However, even in the absence of such compounding factors, urban stormwater drainage remains a degrading influence on stream ecosystems and their water quality, increasing the frequency and magnitude of polluted flows. This is because conventional stormwater drainage (sensu Burns et al., 2012) greatly increases the likelihood that sediments or other materials, and their associated pollutants, on urban surfaces will be carried to the stream when it rains, or that a spill on an impervious surface anywhere in a catchment (e.g. liquid waste poured down a drain, a car washed on a street, firefighting runoff) will flow to a stream even in dry weather. The resulting pulses of polluted flow create frequent hydraulic and chemical disturbances to in-stream communities (Walsh, Roy, et al., 2005). The range of urban activities and materials, together with the efficient hydraulic connection between catchment and stream, result in the delivery of a diverse cocktail of pollutants to stream ecosystems (Kaushal et al., 2020).

Partly in response to a growing awareness of these impacts, alternative drainage approaches designed to reduce the impacts of conventional stormwater drainage have developed and matured since the 1990s (Fletcher et al., 2014). The dominant practice in these alternative approaches was initially largely centralized, end-of-pipe treatment using detention basins, wet and dry ponds or wetlands. This is now shifting to smaller, dispersed systems relying on filtration and biological uptake of pollutants (bio-filtration), with harvesting and export of excess runoff increasingly employed in recent years (Mitchell et al., 2007; Petrucci et al., 2013). We term all of these alternative approaches collectively 'stormwater control.' The evolution of stormwater control has led to a complex mix of technologies employed globally. To add further complexity, motivations for stormwater control extend beyond, and often displace or take priority over, protection of streams. Other motivations include: protection of larger downstream water bodies (coastal or lentic), combined sewer system management to reduce overflow frequency, mitigation of local flooding, and increasingly, the many co-benefits of green infrastructure to city environments (Kim & Song, 2019; Lamond & Everett, 2019).

Such complexity presents challenges for measuring and comparing the effectiveness of large-scale stormwater control on stream ecosystems. Many experimental studies of the hydrologic and chemical performance of individual stormwater control measures (Bettez & Groffman, 2012; SCMs: e.g. Bratieres et al., 2008), and of the cumulative effects of SCMs on the runoff from precincts (up to several ha in area , e.g. Bedan & Clausen, 2009; Wilson et al., 2015), have demonstrated the effectiveness of SCMs in retaining and treating runoff to reduce runoff volume and pollutant loads. Studies of larger-scale cumulative effects of SCMs on stream ecosystems have been less common, with mixed results that are difficult to compare. The difficulty of comparison arises in part from a lack of standard methods for measuring cumulative performance of the diverse forms and scales of SCMs compared to the cumulative impact of urban drainage (Walsh et al., 2022). The relative rarity of such studies is in part a result of the challenges of conducting the necessarily long-term, large-scale experiments required to robustly assess the effects of SCMs on stream ecosystems.

Cumulative effects of dispersed stormwater control on stream water quality have been inferred from correlations with urban density and degree of stormwater control among catchments (Hale et al., 2014; Pennino et al., 2016), or by assessment over time after SCM deployment either in newly developed (Hopkins et al., 2017; Selbig & Bannerman, 2008), or established catchments (Roy et al., 2014) in comparison to control or reference catchments. Urban density has most commonly been measured as total imperviousness, and stormwater control has been measured either categorically (Hopkins et al., 2017; Roy et al., 2014; Selbig & Bannerman, 2008), by proportion of total catchment area upstream of SCMs (Pennino et al., 2016), or by SCM area as a proportion of total catchment area (Hale et al., 2014). Bell et al. (2016) used 'unmitigated' imperviousness: proportional catchment impervious area not draining to an SCM. Other studies have accounted for stormwater control afforded by informal drainage to impervious surfaces either using imperviousness weighted by distance to drain (Walsh & Kunapo, 2009), or by stormwater pipe density (Baruch et al., 2018).

In this study, we used the approach outlined by Walsh et al. (2022) of integrating the degrading effects of stormwater runoff and the restorative effects of SCMs into two comparable measures: EI_{S1}, the proportion of catchment connected to stormwater drainage pipes assuming stormwater control measures have no effect; and EI_S, EI weighted by the stormwater impact index, which measures the cumulative modeled performance of upstream SCMs. Such an approach permits comparison of degradation and restoration trajectories in response to conventional drainage and stormwater control, respectively. It also presents an approach that could be applied across regions and cities for better comparison of studies.

We use these EI measures to assess the effects on stream water quality of the experimental implementation of dispersed SCMs in 6 sub-catchments of 2 independent catchments. The experiment was designed as a before-after-control-reference impact experiment conducted over 19 years, comparing changes in the 6 experimental reaches in the 7 years during and after SCM implementation with changes in the preceding 12 years, and with changes over the full study period in streams of 2 urban control and 3 forested reference catchments. More broadly, the experiment aimed to test if intensive application of dispersed SCMs can restore stream hydrology, water quality and ultimately ecological state (Walsh, Fletcher, et al., 2005; Walsh et al., 2015). In this paper we focus on water quality.

Most studies assessing the water quality effects of SCMs focus on annual loads, primarily to assess export to well buffered large downstream waters. Such long-time-scale measures of water quality are of less direct relevance to smaller, more dynamic stream and river ecosystems (Gomi et al., 2002). We therefore assess the effects on water quality during dominant conditions: in dry weather and following rain events up to 20 mm/d. Such conditions occur >96% of the time in our study streams, and are therefore likely to be prime determinants of ecological structure and function. Dispersed SCMs are typically not designed to retain and treat runoff from storms larger than ~20 mm/d. Rainfall depths of this magnitude are likely to exceed the runoff retention capacity of even natural catchments in this region (Hill et al., 1998), meaning that both urban and non-urban streams will experience hydraulic and water quality disturbance in such rainfalls.

Comprehensively accounting for the multitude of urban contaminants that are responsible for urban stormwater impacts on urban water quality is a near-impossible endeavor (Kaushal et al., 2020). We thus follow the approach implicit in most legislated water quality targets of measuring variables that are nutrients of primary importance to biological processes and in forms with varying degrees of availability; from less bioavailable (i.e. total nitrogen, TN; total phosphorus, TP) to more bioavailable (ammonium, NH_4^+; nitrate + nitrite, NO_x; filterable reactive phosphorus FRP). The variables measured can serve as surrogates for contaminants that are mobile through soils and bioavailable (NO_x) or conservative (electrical conductivity, EC, as a surrogate for dominant ions), or that are bioavailable and readily attach to particles (FRP, NH_4^+). We also measured the particles themselves (TSS), and temperature as a fundamental physical property driving in-stream biological processes.

In this paper, we use hierarchical linear models to assess the effects of the experimental implementation of dispersed SCMs on the selected water quality variables in receiving streams. The experimental design permits us to use the models to make general predictions of the effects of conventional urban stormwater as indicated by EI_{S1}, and of varying density and performance of stormwater control as indicated by EI_S. The predictions will permit better planning and implementation of stormwater control at a catchment-scale for protection of streams and rivers.

2 Materials and Methods

2.1 Study area and experimental design

The study streams are in the Dandenong Ranges on the eastern fringe of the Melbourne metropolitan area, Victoria, in temperate south-eastern Australia (see Walsh et al. 2021). The study was a before-after-control-reference-impact experiment, with 7 sites in independent catchments comprising: three reference, forested catchments (Sa, Ly, Ol), with little or no stormwater drainage infrastructure; two control urban catchments (Br, Fe), with streams degraded by urban stormwater drainage, and two experimental urban catchments (L4 and D4), in which stormwater control measures were implemented progressively from 2009 and 2012, respectively. These sites were a subset of the sites studied by Hatt et al. (2004), selected to minimize variation in physiographic and climatic conditions of their catchments. These seven independent sites were sampled from 2001 to 2019. From November 2010, four additional experimental sites (L1, Ln, and Ls, all upstream of L4 and D8, downstream of D4) were added to the sampling program. Stormwater control measures were implemented in the larger D8 catchment from 2011.

The determination of temporal changes in EI resulting from urban growth and from SCM implementation in the experimental catchments, was described in detail by Walsh et al. 2021. Briefly, in 2001, EI ranged from 0 to 1% in the reference catchments, and from 2 to 25% the experimental and control catchments (Table 1). Over the 19-year study period, EI grew by 0-1% in the reference sites (e.g. Sa grew from 1.04% to 1.05%) and by 4-14% in the experimental and control catchments (2001 vs 2019 $\mathrm{EI_{S1}}$ in Table 1). Implementation of 638 SCM projects (stormwater harvesting tanks and raingardens, at scales from residential property to sub-catchment of council stormwater pipes) in the 6 experimental catchments reduced EI by 13-68% (2019 $\mathrm{EI_S}$ vs $\mathrm{EI_{S1}}$ in Table 1).

The experimental design thus permitted a comparison of temporal trends in water quality variables in the 6 experimental sites after commencement of SCM implementation compared to temporal trends in those sites in the period before implementation, and temporal trends through both the before and after periods for the reference and control streams.

Table 1. Catchment statistics and experimental class (reference, control or experimental) for the 11 study sites. $\mathbf{EI_{S1}}$ is the proportion of catchment connected to stormwater drainage pipes assuming stormwater control measures have no effect; and $\mathbf{EI_S}$ is EI weighted by the stormwater impact index, which measures the cumulative modeled performance of upstream SCMs. @ >p(- 14) * >p(- 14) * >p(- 14) * >p(- 14) * >p(- 14) * >p(- 14) * >p(- 14) * >p(- 14) * @

Code

&

Stream

&

Class

&

Catchment area (km^2)

&

EI_{S1} (2001)

&

EI_{S1} (2019)

&

EI_S (2019)

&

Septic tank density (N/km^2)

Ly

&

Lyrebird

&

Reference

&

7.2

&

0.0

&

0.0

&

0.0

&

0.1

Ol

&

Olinda

&

Reference

&

9.1

&

0.1

&

0.1

&

0.1

&

75.5

Sa

&

Sassafras

&

Reference

&

1.9

&

1.0

&

1.1

&

1.1

&

112.8

Br

&

Brushy

&

Control

&

14.9

&

21.7

&

24.5

&

24.5

&

9.9

Fe

&

Ferny

&

Control

&

6.4

&

11.2

&

11.9

&

11.9

&

44.4

L4

&

Little Stringybark

&

Experimental

&

4.5

&

8.9

&

10.2

&

6.4

&

17.8

Ls

&

Little Stringybark Sth

&

Experimental

&

1.0

&

11.8

&

13.4

&

6.0

&

21.7

L1

&

Little Stringybark Central

&

Experimental

&

0.8

&

22.4

&

25.5

&

22.2

&

3.0

Ln

&

Little Stringybark Nth

&

Experimental

&

1.5

&

4.9

&

6.2

&

2.0

&

22.7

D4

&

Dobsons

&

Experimental

&

3.5

&

2.5

&

2.5

&

2.0

&

50.8

D8

&

Dobsons downstream

&

Experimental

&

7.9

&

1.9

&

2.0

&

1.4

&

44.3

*Ls, L1 and Ln are independent tributaries upstream of L4, † D4 is upstream of D8

2.2 Data collection

Water samples were collected from the seven sites in independent catchments from 2001 to 2019 over three periods. Regular samples were collected every second week for 29 months from Sep 2001 to Jan 2003, and monthly for 21 months from Feb 2004 to Nov 2005 (except for Br and Fe) and for 10 years from May 2009 to Jul 2019. In addition to the regular samples in these periods, we collected samples during 6–12 rain events per year. From Nov 2010, the four additional experimental sites were also included in the sampling program. Sites were sampled in one of four possible orders on each sampling occasion to ensure samples from each site were taken at a range of times.

We recorded EC with a TPS Direct Reading Conductivity Meter Model 2100 (https://tps.com.au/) and temperature with a 'H_2O' water quality multiprobe (https://www.hydrolab.com/) in 2001-2002, and both variables using a Horiba U-10 multiprobe (https://www.horiba.com/) or a YSI 6920 V2 multiprobe (https://www.ysi.com/) thereafter. We collected samples for TN, TP, NO_x, NH_4^+, filterable reactive phosphorus (FRP) and total suspended solids (TSS). Samples for NO_x, NH_4^+, and FRP were filtered in the field using 0.2 m filters from 2001 to 2003, and 0.45 m filters thereafter. This change of filter type was inconsequential: see consideration of the *filter* predictor below. All bottles were stored on ice until returned to the laboratory (< 5 h). All analyses were performed in a NATA (National Association of Testing Authorities, https://nata.com.au/) accredited laboratory following standard methods (APHA et al., 2012).

Very few records were below detection limit. Two FRP values (<0.1% of all FRP records), 8 NO_x values (0.4%), and 11 NH_4^+ values (0.5%) were less than the detection limit of 0.001 mg/L. No TN values were less than the detection limit of 0.2 mg/L, 19 TP values (0.8%) were less than the detection limit of 0.1 mg/L and 1 TSS value (<0.1%) was less than the detection limit of 0.5 mg/L. For all variables, below-detection-limit records were set at half the detection limit in the statistical models.

Determination of impervious areas, drainage connection, and SCM implementation and performance over the study period, used to calculate EI-related predictor variables, was described in detail by Walsh et al. (2021). Rainfall predictor variables were calculated from 6-minute-time-step rainfall data over the study period. For each catchment, a weighted average 6-minute rainfall time series was estimated using available gauge data (Bureau of Meteorology's Montrose (086076) and Ferny Creek (86266) gauges; Melbourne Water's Brushy Creek (229249A), Dandenong Creek (228373A), Silvan (586177), Mt Dandenong (586090), and Mt Evelyn (229690A) gauges, and five gauges installed by the project team over different periods in the L4 and D8 catchments), and the Bureau of Meteorology daily rainfall grid (See Walsh et al 2021, Appendix S3 for a detailed description of the methods used to calculate catchment-weighted averages). Septic tank density was estimated from a dataset of unsewered properties supplied by Yarra Ranges Council in 2000, and South-East Water for the D4 catchment in 2016. These data are likely to be reliable for the study period as there were no sewer upgrade programs in the study catchments.

2.3 Statistical analyses

For each of the eight water quality variables (FRP, TP, NO_x, NH_4^+, TN, TSS, EC and temperature) we assessed the effect of dispersed stormwater control measures using a hierarchical linear model as described by Walsh et al. (2022).

The two primary stormwater-related effects of the model are:

- *degrd*, the putative degrading effect of urban stormwater runoff, represented by $\log_{10}(EI_{S1} + 0.1)$, where EI_{S1} is percentage effective imperviousness assuming stormwater control measures have no effect;
- *restr* the putative restorative effect of SCMs, represented by $\Delta \mathrm{E}I_S$, the difference between $\log_{10}(EI_{S1} + 0.1)$ and $\log_{10}(EI_S + 0.1)$, where EI_S is effective imperviousness, with impervious areas upstream of stormwater control measures weighted by S, a metric of performance. S measures reduction in uncontrolled runoff frequency, runoff volume, and contaminant concentrations, and restoration of filtered flows: see Walsh et al. (2022)

In the before phase of the experiment (2001 to 2009), the six experimental sites had EI_{S1} values spanning the range of EI_{S1} between the control sites and reference sites. In all but two sites (the reference sites Ly and Ol, Table 1), EI_{S1} increased over the study period, but this growth in imperviousness was small in the log domain (Figure 1). $\Delta \mathrm{E}I_S$ was zero for the entire study period for the reference and control sites, and for the experimental sites in the before phase. It became increasingly negative after the stormwater control measures began being implemented. If the stormwater control measures completely mitigated the effect of urban stormwater runoff, then the *degrd* (EI_{S1})and *restr* ($\Delta \mathrm{E}I_S$) effect sizes should be equivalent.

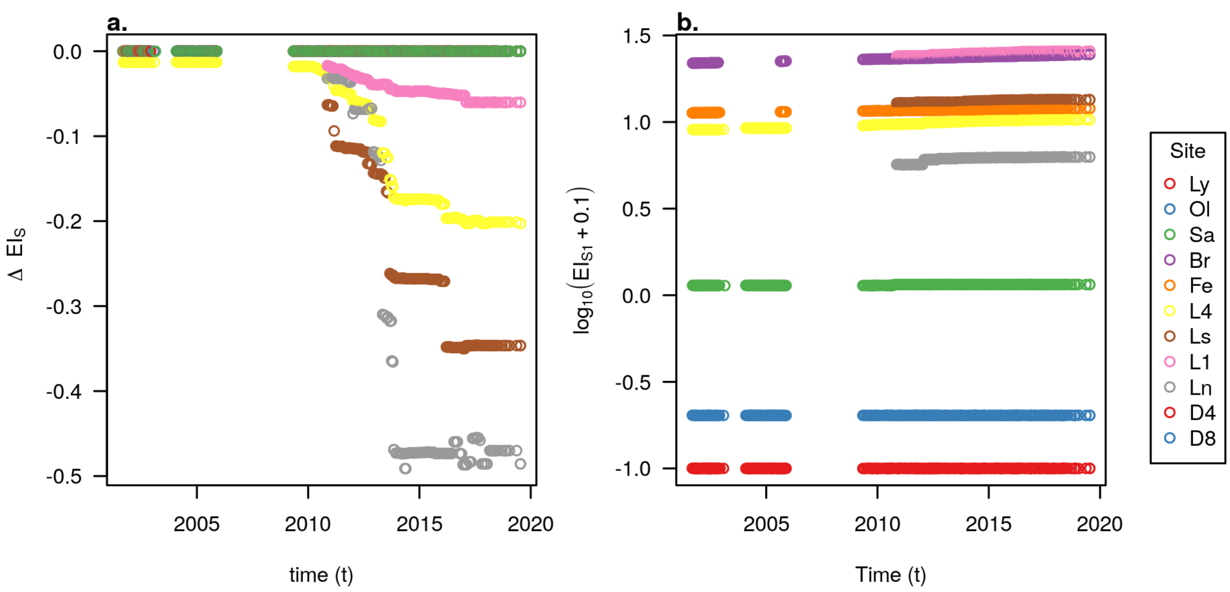


Figure 1. Variation in a. restr ($\mathbf{\Delta EI_S}$) and b. degrd ($\mathbf{\log_{10}(EI_{S1} + 0.1)}$) in each site over the study period for every water quality sample used in the models. All eight water quality response variables were modeled as being drawn from a normal distribution:

$$y_i = \text{Normal}(\mu_i, \sigma)$$

(eq. 1)

where y_i is the value of the response variable in the ith sample, and μ_i is the mean estimate for the sample and σ is the residual standard deviation. To approximate such a distribution, FRP, TP, NH_4^+, TN, TSS and temperature were $\log_{10}$-transformed, and NO_x and EC were square-root transformed.

The basic model assessing the effects of *degrd* and *restr* in the before-after-control-reference-impact experiment (Walsh et al., 2022) estimated μ_i as:

$$\mu_i = \alpha + \alpha[\text{site}_j] + \beta_D \text{degrd}_i + \beta_R \text{restr}_i + \beta_T[\text{site}_j]t_i + \beta_A \text{auto}T_i$$

(eq. 2)

where α is the global intercept; $\alpha[\text{site}_j]$ is the random variation to that intercept for site *j*; β_D represents the effect of *degrd*; β_R represents the effect of *restr* (only non-zero in experimental sites after SCM manipulation had begun); $\beta_T[\text{site}_j]$ represents the effect of time, t, within the site *j*; β_A represents the effect of temporal autocorrelation, *autoT*, among samples from each site. We thus modeled a random effect of site in the intercept and in the slopes of the time effects (allowing for different trends over time among sites unrelated to the experimental effect), but fixed effects of *degrd* and *restr* . However, models of temperature, NO_x and TN did not consistently converge with a variable time effect, and as the time effect was near-identical among sites, we modeled time as a fixed effect for these variables.

For each model, *autoT* was the mean residual value from a model with the same structure, for the preceding 45 days in each site. The use of residuals of models without autocorrelation terms has been demonstrated as an effective method of accounting for autocorrelational effects (Crase et al., 2012). In a preliminary analysis we compared model fits using *autoT* calculated using antecedent periods from 1 to 120 days, and found 45 days to be the optimal period for this dataset.

Stream contaminant concentrations and temperature vary in response to flow (e.g. Guo et al., 2020), and both urban stormwater drainage and SCMs alter the response of stream flow to rainfall. We thus used antecedent rainfall as a predictor variable in our models to avoid conflation of the effect of flow on stream concentrations and the effect of urban stormwater runoff and SCMs on stream flow. As stormwater drainage conveys impervious runoff rapidly to the stream, recent rainfall (within hours) is a more likely predictor of change to flow than less recent rainfall. Similarly, the potential for stormwater control measures to have sufficient void to prevent rapid runoff is likely to be predicted by recent rainfall. We thus included rainfall depth in the preceding 24 h (*rain1*) of each sample as a fixed predictor, as well as the interactions of *rain1* with *degrd*, and with *restr*, on the premise that the effects of stormwater drainage and SCMs will likely vary with depth of rainfall events.

Our study design included sites with spatial dependence: L1, Ln, and Ls are 1.7, 1.7 and 1.4 stream-km upstream of L4, respectively, and D4 is 0.9 stream-km upstream of D8. We thus also included a spatial autocorrelation term, *autoS*, in all models, which was applied to samples from L4 and D8. *autoS* was the mean of values recorded in L1, Ln, and Ls on the same date for L4, and the value recorded in D4 on the same date for D8.

The primary model of each variable (for which the time effect varied among sites) was thus:

$$\mu_i = \alpha + \alpha[\mathrm{site}_j] + \beta_D \mathrm{degrd}_i + \beta_R \mathrm{restr}_i + \beta_p * \mathrm{rain1}_i + \beta_{\mathrm{pd}}(\mathrm{rain1}_i * \mathrm{degrd}_i) + \beta_{\mathrm{pr}}(\mathrm{rain1}_i * \mathrm{restr}_i) + \beta_t[\mathrm{site}_j] t_i + \beta_{\mathrm{at}} \mathrm{auto}T_i$$

(eq. 3)

We aimed to ensure that each model adequately accounted for sources of variation unrelated to our experimental manipulation. We therefore calculated and compared the primary model to a range of more complex models. For each variable, we added up to three additional predictors to the primary model, and improvement in model fit was assessed by differences in the leave-one-out estimate of out-of-sample predictive fit ($\mathrm{ELPD}_{\mathrm{loo}}$: Vehtari et al. (2017)) to select the best-fit model for each response variable. Additional variables included in candidate models were:

- *season*, a sinusoidal curve with a period of 1 year, a maximum of 1 on the summer solstice (December 21) and a minimum of -1 on the winter solstice. This variable represents seasonal changes in climate and instream biological activity, and was included in models for all variables as an additional term $\mathrm{seas} * \mathrm{season}$
- *rain365*, sum of antecedent rainfall depth over the preceding year. This variable is an indicator of longer-term water storage and baseflow contribution from catchment soils, and was included ($\log_{10}$-transformed) in models for all variables as an additional term $\beta r365 * rain365$.
- *channel*, which was zero for all samples, except for samples from L4 and Ln for two years after 300 m of vegetated channel and banks, ~ 300 m upstream of Ln, were disturbed by mechanical re-profiling and vegetation removal, resulting in liberation of fine sediments. The variable represents the potential for the exposed sediments to mobilize contaminants during the period that channel vegetation returned. We included this variable in candidate models of FRP, TP, NO_x, NH_4^+, TN, and TSS as an additional term $\mathrm{c} * \mathrm{channel}$.
- *diel*, a sinusoidal curve with a period of 1 day, a maximum of 1 at 1800 h, when stream temperature is likely to be at its daily maximum, and

minimum of -1 as 0600 h. This variable was only applied to models of temperature as an additional term $\beta_{\text{diel}} * \text{diel}$. All candidate temperature models included both *season* and *diel*, as we expected such temporal variation was likely to be a dominant determinant of stream temperature. We also considered a model with interactions of season with *restr* and *degrd* ($\text{sea}s_d * \text{season} * \text{degrd} + \text{sea}s_r * \text{season} * \text{restr}$), to represent the possibility that stormwater effects on temperature vary with season; and a model with those interactions together with interactions of diel with *restr* and *degrd* ($\text{diel}_d * \text{diel} * \text{degrd} + \text{diel}_r * \text{diel} * \text{restr}$), to represent stormwater effects on diel variation in temperature.

- *septic,* the density of septic tanks in each catchment. Because of the mobility of NO_x through soils, and the variability in sewerage infrastructure among the study catchments, NO_x and TN concentrations in the study streams are likely to be more strongly predicted by *septic* than by effective imperviousness (Hatt et al., 2004). We included this variable (square-root-transformed) in all candidate models of NO_x and TN ($\text{e} * \text{septic}$), and also trialed the addition of its interaction with *rain1* ($\text{pe} * rain1 * \text{septic}$).

- *filter,* a binary variable distinguishing samples of FRP, NO_x and NH_4^+ taken up until 2003, using 0.2- m filters from those taken later using 0.45- m filters, to account for any temporal variation in those variables that may have resulted from this methodological change. We assessed the effect of including *filter* in all candidate models (see below) of FRP, NO_x and NH_4^+ ($\text{f} * \text{filter}$). The *filter* effect for FRP and NH_4^+ was near-zero, and its inclusion in the NO_x model introduced multicollinearity with the effect of time. The inclusion or exclusion of a filter effect did not change the estimates of the experimental effects in any of the models (See Text S1, Figure S2). We thus concluded that the effect of changing filters after 2003 was inconsequential and elected to not consider *filter* further, or include it in candidate models reported here.

Combining the additional variables (other than *filter*) with the primary model (eq. 3) resulted in comparison of 8 models for FRP, TP, NO_x, NH_4^+, TN, and TSS (primary model plus combinations of *channel*, *season*, *rain365*), 16 models for NO_x and TN (additional 8 models including *septic* with and without an interaction with *rain1*), 4 models for EC (primary model plus combinations of *season* and *rain365*), and 6 models for temperature (primary model plus *season* and *diel* and combinations of *rain365* and interactions as described above). See Appendix S1: Table S1, for the full list of models.

All α and β parameters were drawn from a weakly informative normal distribution (mean 0, standard deviation 5), except for the random site parameters s and t, which were drawn from hyperdistributions with a mean drawn from a normal distribution with mean 0 and standard deviation 5, and standard deviation drawn from a half-Cauchy distribution (mean 0, standard deviation 2). We derived the models using the Markov Chain Monte-Carlo sampler of Stan (Carpenter et al., 2017), ensuring standard diagnostic tests of model performance were satisfied, and that each model provided accurate predictions of the data. See Text S1 for further details.

2.3.1 Model prediction for assessment of stormwater control effects

To aid interpretation of the predictions of water quality responses to the degrading effects of urban stormwater (*degrd*) and the potentially restorative effects of SCMs (*restr*), we made general predictions of each response variable to a range of predictor variables by setting site-specific random parameters to their mean values.

We first explored the primary effects of interest (*restr*, *degrd* and *rain1*) and their interactions, first by comparing the response of each variable to rain1 under three scenarios:

1. a stream with EI_{S1} equivalent to each experimental site at the end of the study, with no stormwater control measures (i.e. each site as it would be without SCM implementation);
2. a stream as in a), but with the stormwater control measures achieved in that catchment at the end of the study (i.e. each site as it was after SCM implementation);
3. a reference stream (zero EI_{S1}, zero SCMs).

For all scenarios for all variables, where relevant, *season* was set to the equinox, except for temperature, for which season was set to the summer solstice and *diel* to 1800 h (i.e. their maximum values); *rain365* to its mean value; *channel* to zero; *septic* to the value for each experimental catchment; and *time* to the end of the study. *autoT* and *autoS* were excluded from the predictions to new data, as they were for modeled data points that had no antecedent or upstream dependent data.

In those contexts we predicted responses to combinations of a range of *degrd* values encompassing and including the 2019 values for each site; a range of *restr* values encompassing and including the minimum values (2019: i.e. maximum SCM implementation) achieved in each site; a range of rain1 values, including 0, 2 and 8 mm (the 50th, 75th, and 90th percentile rain1 values for all days between 2001 and 2019) and 20 mm (the typical maximum storage of installed SCMs). They thus represent dominant conditions. The 75th percentile is of particular importance because regional water quality concentration objectives for TP, TN and EC are set to the 75th percentile concentration (EPA Victoria, 2021). To assess the potential for SCM implementation to meet these management targets, we compared predicted concentrations of these variables for the 75th percentile rainfall to the relevant objectives for our study streams (Central foothills and coastal plains, Yarra lowlands).

For *rain1* = 0, 2, 8 and 20 mm, we calculated differences in posterior distributions of each response variable between scenarios:

- a) minus b) (above) to assess the degree to which SCM implementation changed the variable;
- b) minus c), to assess the degree to which the response variable approached reference condition after SCM implementation.

To compare the response of each variable to *restr* (i.e. SCM implementation) with that to *degrd* (i.e. inferred impact of stormwater drainage from EI), we first plotted the predicted response to *degrd* (EI_{S1}) for *rain1* = 0, 2, 8 and 20 mm, assuming no SCMs (i.e. *restr* = 0) and setting other predictors as described above. For each experimental site, we plotted two points on each of these plots, corresponding to scenarios a) and b) above, to indicate the direction of response to SCM implementation. The EI value assigned to scenario a) equaled EI_{S1} and to b) equaled EI_S.

3 Results

3.1 Overview

SCMs, as indicated by *restr*, reduced TP, FRP, TN, NOx and NH4+ concentrations and temperature, and increased EC; most strongly in dry weather, with reduced effect sizes following increasingly large rain events. For TP, FRP, and NH4+, responses to *restr* after zero to ~8 mm of rain were similar to responses to *degrd*: i.e. the restorative effect of SCMs reversed the degrading effect of stormwater drainage. TN and NO_x concentrations, among the study sites were most strongly explained by septic tank density: the reduction of these contaminants by SCMs is likely a result of reductions of stormwater runoff volume, and removal of N in stormwater runoff that may receive some runoff from septic systems. Increased EC (adding to the effect of *degrd* rather than reducing it) and reduced temperatures suggest SCMs increased groundwater flows into the streams. TSS, which was lower in control and experimental streams than in reference streams during low flows, was not affected by SCMs.

Before describing these primary results in detail, we describe the results of model selection, and the effects of variables other than the primary variables of interest (*restr*, *degrd*, and *rain1*).

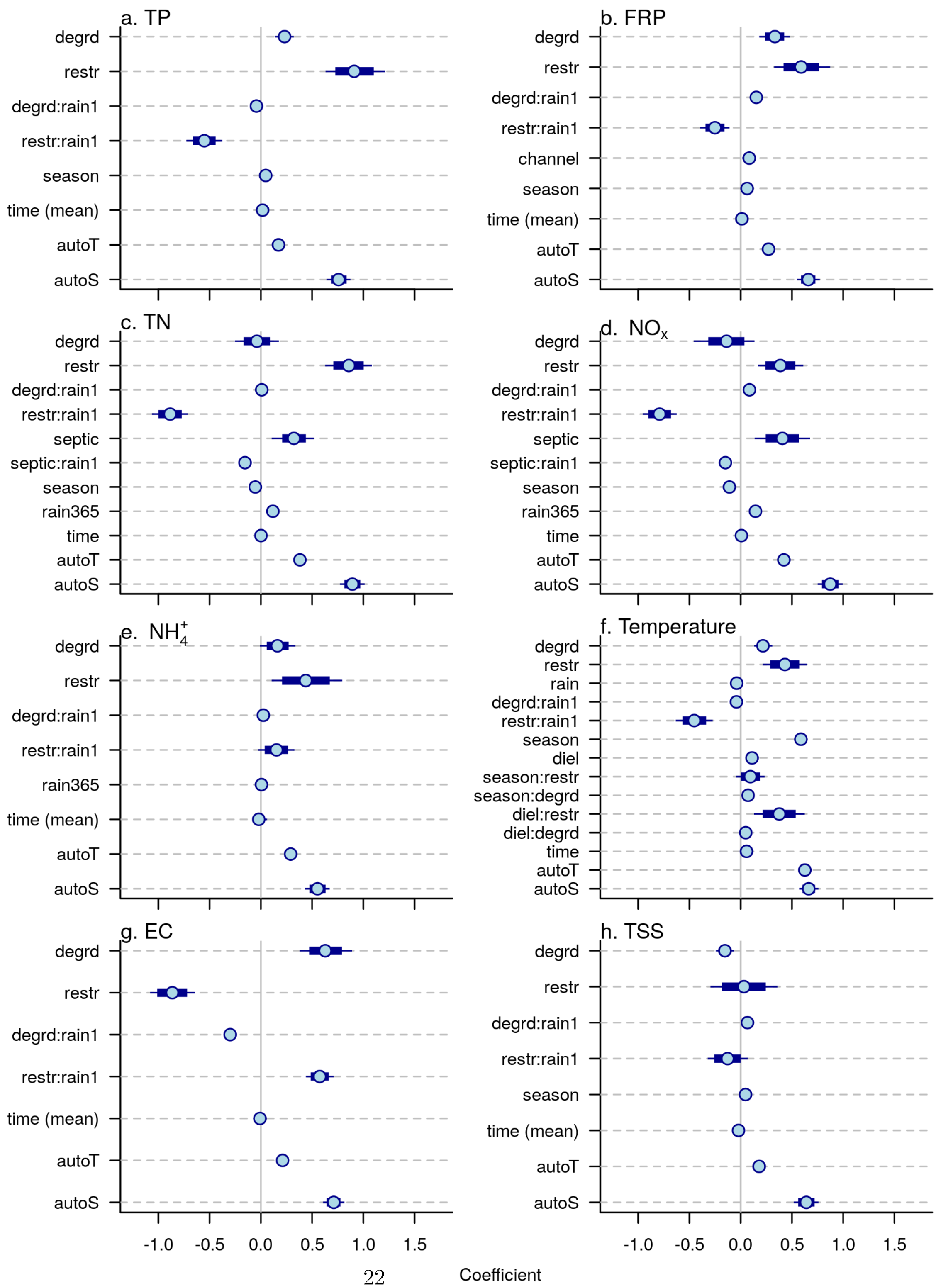
3.2 Model selection and assessment

For each response variable, the model with maximum $ELPD_{loo}$ was selected for assessment (Text S1: Table S1). A clearly best model was evident for NO_x, TN and temperature (Table S1). For other variables, addition of *season*, *rain365* and *channel* to the primary model (Eq. 3) made only small differences to model fit. Additional variables made no consequential difference to the effects of the predictors of primary interest (*restr, degrd, rain1* and interactions). Models selected for assessment were:

- FRP, primary model + *channel* + *season*;

- TP and TSS, primary model + *season*;
- NH_4^+, primary model + *season* + *rain365*;
- EC, primary model;
- NO_x and TN, primary model (with fixed time effect) + *septic* + *septic:rain1* +*season* + *rain365*;
- Temperature, primary model (with fixed time effect) + *season* + *diel* + *rain365* + *season:degrd* + *diel:degrd*+ *season:restr* + *diel:restr*.

All eight selected models predicted their response variable well (R of predicted to observed 0.76-0.94, Figures S3-S10). In all models, the *autoT* and *autoS* effects were strongly positive (Figure 2), and the addition of these autocorrelation terms improved the estimation of other parameters without significant change to mean estimates. The effect of *time* did not differ from zero for any response variable except temperature (Figure 2f), for which an increase of 0.9°C over the study period was predicted. This estimate closely matches the increase in mean air temperature of 0.95°C over the same period in a nearby weather station (Scoresby, <http://bom.gov.au>). *Rain365* had a positive effect on TN and NO_x (Figure 2c, d, e) suggesting their concentrations increased in wetter periods with higher baseflow contribution to streams. *Season* had a small positive effect on TP, FRP and TSS (Figure 2a, b, h), meaning their concentrations tended to be higher in summer: conversely TN and NO_x were lower in summer (Figure 2c, d). *Season* and *diel* both had positive effects on temperature, pointing to highest temperatures in summer and late afternoon, but these effects interacted with the *restr* and *degrd*, because urban stormwater impacts increase the amplitude of seasonal and diel temperature changes. In assessing temperature effects below, we focus on periods of maximum temperature, as these are likely to be of greatest significance to biotic response. The channel works upstream of Ln in 2016 (*channel* effect) increased FRP by 18.9% (95% credible interval 4–35.3%) (Figure 2b).



a. TP
degrd
restr
degrd:rain1
restr:rain1
season
time (mean)
autoT
autoS
b. FRP
degrd
restr
degrd:rain1
restr:rain1
channel
season
time (mean)
autoT
autoS
c. TN
degrd
restr
degrd:rain1
restr:rain1
septic
septic:rain1
season
rain365
time
autoT
autoS
d. NO_x
degrd
restr
degrd:rain1
restr:rain1
septic
septic:rain1
season
rain365
time
autoT
autoS
e. NH_4^+
degrd
restr
degrd:rain1
restr:rain1
rain365
time (mean)
autoT
autoS
f. Temperature
degrd
restr
rain
degrd:rain1
restr:rain1
season
diel
season:restr
season:degrd
diel:restr
diel:degrd
time
autoT
autoS
g. EC
degrd
restr
degrd:rain1
restr:rain1
time (mean)
autoT
autoS
h. TSS
degrd
restr
degrd:rain1
restr:rain1
season
time (mean)
autoT
autoS
-1.0
-0.5
0.0
0.5
1.0
1.5
Coefficient

Figure 2. Coefficient plots for models of a. total phosphorus, b. filterable reactive phosphorus, c. total nitrogen, d. nitrate + nitrite, e. ammonium, f. temperature, g. electrical conductivity, h. total suspended solids. Each plot shows the mean, 80% (thick line) and 95% (thin line) confidence intervals for the fixed effects (see text for effect definitions), and the mean time effect in models in which time varied by site (see Figures. S3-S10 for site-specific random effects). The coefficient axes are on a common scale as all response variables were centered and standardized.

3.3 Degradation and restoration effects interact with antecedent rain

The effects of primary interest—*degrd* and *restr*, and *septic* for TN and NO_x—varied in their effect among response variables, and in most cases these effects varied with *rain1* (interaction terms in Figure 1). The form of these interactions is illustrated for Ln (the catchment with the greatest stormwater control) in Figure 3, and differences between scenarios a) and b), and b) and c) under the 4 rain conditions are shown for all 6 experimental sites in Figure 4.

TP concentrations increased strongly with *rain1* in all study streams, but were consistently lower in reference streams than in control streams. SCMs in Ln reduced TP concentrations to reference levels after little or no rain (Figure 3a, 4a.II). As a result, Ln met the Victorian government objective for TP concentration after SCM implementation, when it would have failed without SCMs (Figure 3A). SCMs also reduced FRP concentrations in Ln after little or no rain, but not quite to reference levels (Figure 3b, 4b.II). FRP concentrations in reference streams increased only weakly with *rain1* up to 20 mm (Figure 3b), so that differences in concentrations between streams with and without urban drainage increased with rainfall. The absolute reduction in both TP and FRP concentrations by SCMs in Ln was similar across all *rain1* values, with increasing uncertainty after larger rain events (Figure 4a.i, b.i), however, log-transformed concentrations after higher rainfall became more like urban control streams than reference streams (Figure 3a, b). SCM-induced reductions in P were observed in Ln, Ls and L4, but reductions in D8, D4 and L1, which received less reduction in EI, were near zero (Figure 4a.i, b.i).

TN, NO_x and NH_4^+ concentrations in control and reference streams were poorly predicted by *rain1* (Figure 3c, d, e). SCMs reduced TN and NO_x after little or no rain in Ln, Ls, L4, D8, and D4, but no reduction was predicted in TN after 8 mm (Figure 4c.I) or in NO_x after 2 mm (Figure 4d.i), and an increase in both TN and NOx was predicted after higher rainfall. The SCM-induced reduction in TN increased the likelihood that Ln would meet the government objective for TN concentration (Figure 3c). These variations in TN and NO_x in response to *rain1* and to SCM installation in control and experimental streams fell within the range of concentrations for reference streams (Figure 3c, d; 4c.ii, d.ii). SCMs reduced NH_4^+ concentrations across the range of *rain1* (Figure 4e.i), making concentrations more consistent with reference concentrations (Figure 3e).

Summer control and experimental stream temperatures were 4-5 C° warmer than reference streams (Figure 3f). SCMs reduced summer stream temperatures in all experimental streams after 0-8 mm of rain (Figure 4f.i). The reduction in temperature in Ln was large enough to approximate reference temperatures after 0 and 2 mm of rain (Figure 4f.ii).

EC was poorly predicted by *rain1* in reference streams, rarely exceeding 0.2 mS/cm, while in control streams EC was typically 0.4-0.7 mS/cm during dry weather, reducing to 0.2-0.4 mS/cm after 20 mm of rain (Figure 3g). SCMs increased EC in all experimental streams during dry weather (Figure 4g.i), increasing the difference from reference condition (Figure 3g, Figure 4g.ii). Ln exceeded the government objective for EC with and without SCM implementation (Figure 3g).

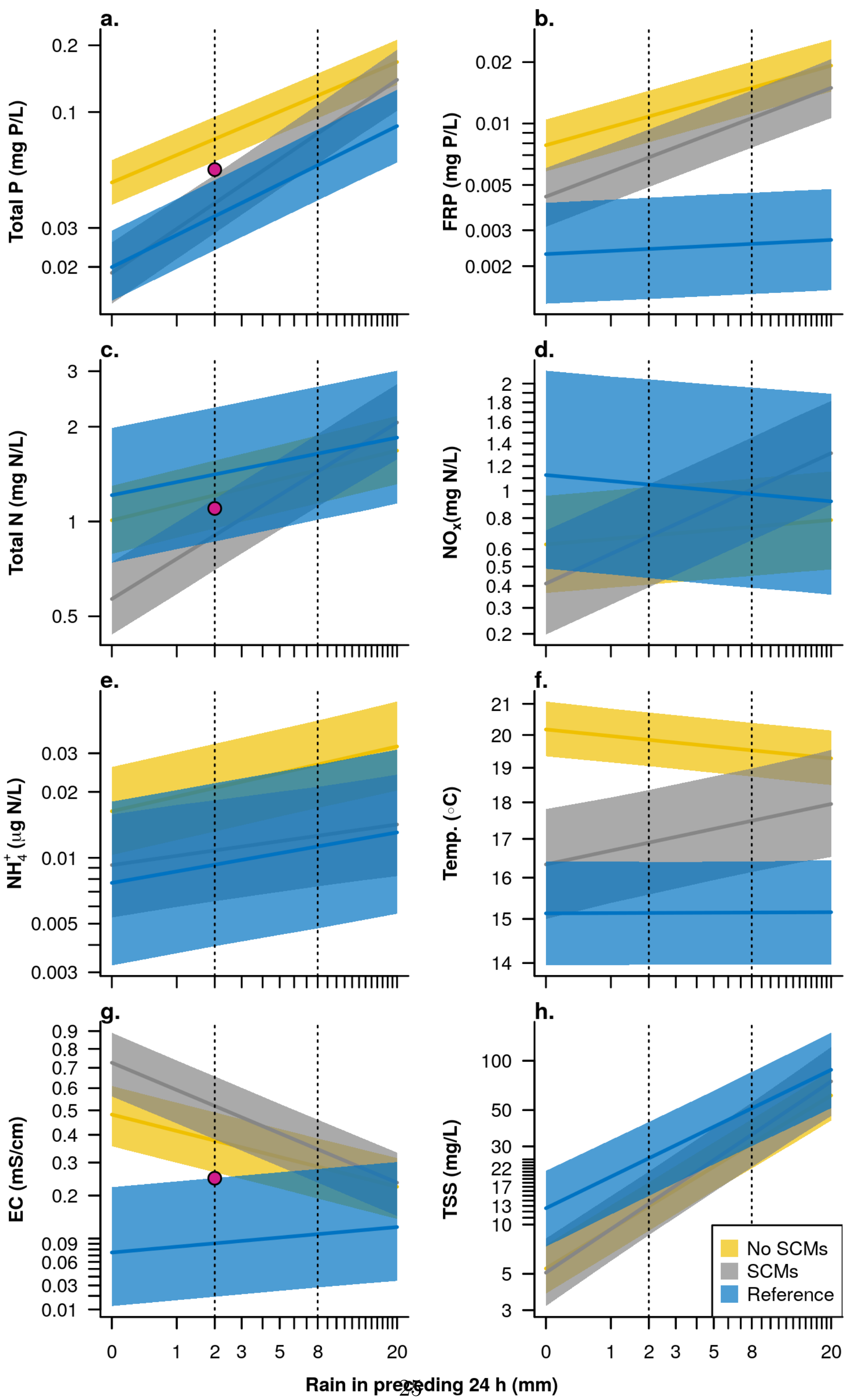
TSS concentrations increased with *rain1*, and reference streams had higher concentrations than control streams after 0-2 mm of rain (Figure2h, Figure 4h.ii). SCMs had no effect on TSS concentrations (Figure 3h, Figure 4h.i).

3.4 Restoration response v. degradation response

Only TP, FRP, temperature and EC showed a strong positive response to EI_{S1} among the study sites (Figure 5a, b, f, g). SCMs reduced TP and FRP concentrations more than predicted by the degradation trend of EI_{S1} in dry weather (Figure 5a.i, b.i). This was also the case for temperature after up to 8 mm of rain (Figure 5f). After 2-8 mm of rain, the response of TP and FRP to SCMs matched the degradation trend closely (Figure 5a.ii, iii, 4B.ii, iii), as did the response of temperature after 20 mm (Figure 5f.iv). After 20 mm, the response of TP and FRP to SCMs was less than predicted by the degradation trend (Figure 5a.iv, b.iv).

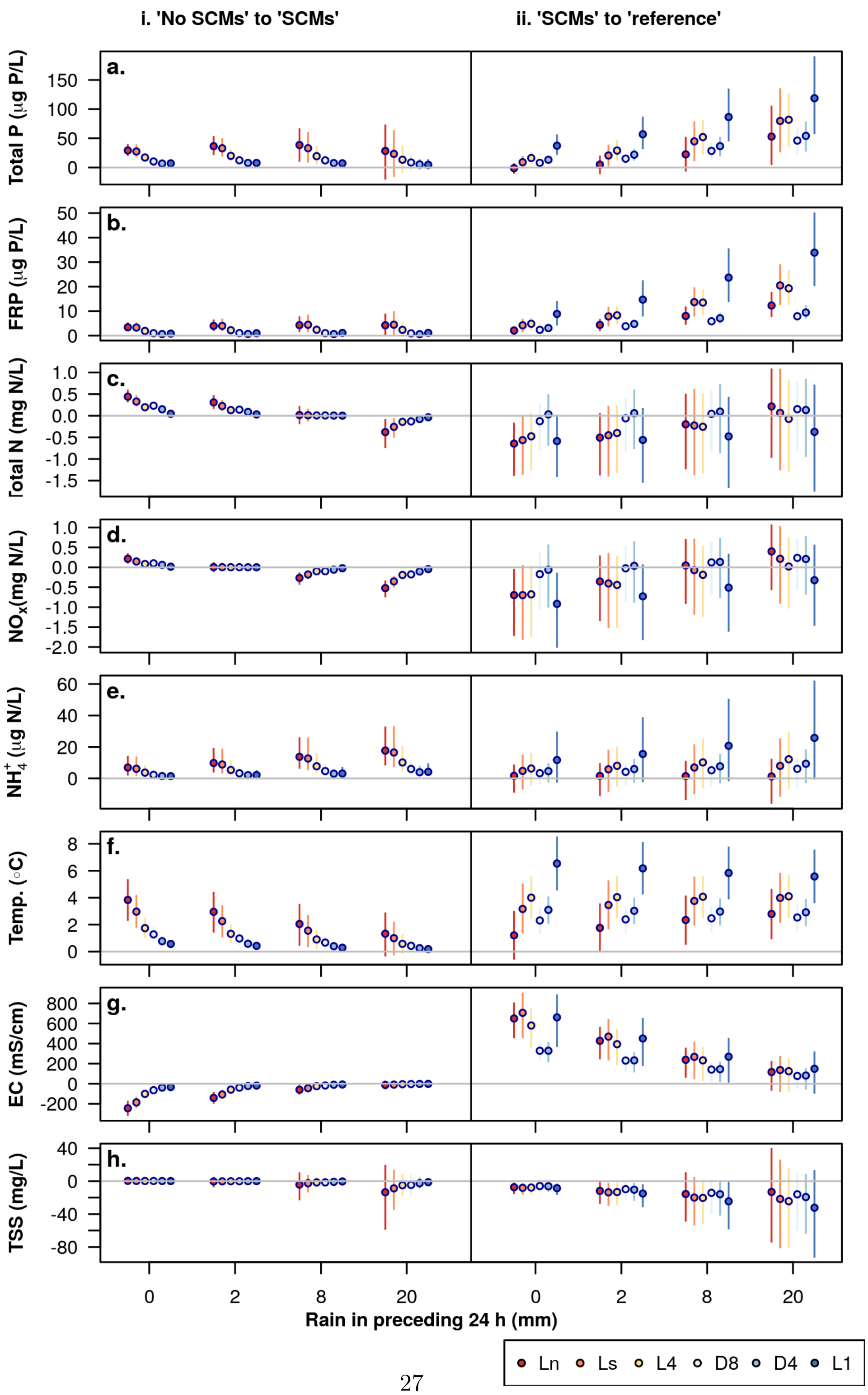
The increase in EC in response to SCMs was in the opposite direction to the degradation trend (Figure 5g). The lack of response of TSS to SCMs was consistent with the EI being a poor predictor of TSS concentrations (Figure 5h).

Although TN, NO_x and NH_4^+ were not well predicted by EI_{S1}, the reduction of EI by SCMs resulted in a reduction in their concentrations in dry weather, and for TN and NH_4^+ after 2 mm of rain, and for NH_4^+ after 8-20 mm (Figure 5c, d, e).



a.
Total P (mg P/L)
0.2
0.1
0.03
0.02
b.
FRP (mg P/L)
0.02
0.01
0.005
0.003
0.002
c.
Total N (mg N/L)
3
2
1
0.5
d.
NOx (mg N/L)
2
1.8
1.6
1.4
1.2
1
0.8
0.6
0.5
0.4
0.3
0.2
e.
NH4+ (µg N/L)
0.03
0.02
0.01
0.005
0.003
f.
Temp. (°C)
21
20
19
18
17
16
15
14
g.
EC (mS/cm)
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.09
0.06
0.03
0.01
h.
TSS (mg/L)
100
50
30
22
17
13
10
5
3
No SCMs
SCMs
Reference
0
1
2
3
5
8
20
Rain in preceding 24 h (mm)

Figure 3. a. Total phosphorus, b. filterable reactive phosphorus, c. total nitrogen, d. nitrate + nitrite, e. ammonium, f. temperature, g. electrical conductivity and h. total suspended solids, predicted as a function of rain1 (rain in the 24 h before sampling) under three scenarios in Ln, the catchment with the greatest stormwater control. In each panel, the lines and polygons show medians and 95% credible intervals, respectively, for scenarios: no stormwater control measures (SCMs) were installed (yellow), SCMs installed by the end of the study in the Ln catchment (grey), reference condition (0% EIS1). The vertical lines indicate the 75th (2 mm) and 90th (8 mm) percentile antecedent 24-h rainfall; the median was 0 mm. The red points in a, c, and g indicate the Victorian government objective for 75th percentile concentration of TP, TN, and EC respectively for Yarra region lowland streams (EPA Victoria, 2021).



i. 'No SCMs' to 'SCMs'
ii. 'SCMs' to 'reference'
a.
Total P (µg P/L)
b.
FRP (µg P/L)
c.
Total N (mg N/L)
d.
NO_x(mg N/L)
e.
NH_4^+ (µg N/L)
f.
Temp. (°C)
g.
EC (mS/cm)
h.
TSS (mg/L)
0
2
8
20
Rain in preceding 24 h (mm)
Ln
Ls
L4
D8
D4
L1

Figure 4. Differences (median and 95% credible intervals) in response variables in each of the six experimental site between three scenarios. The left panels show the difference between the response variable without SCMs ("No SCMs") and with SCMs as installed at the end of the study in each site ("SCMs"). The right panel shows the difference between the response variable with SCMs and reference condition. a. total phosphorus, b. filterable reactive phosphorus, c. total nitrogen, d. nitrate + nitrite, e ammonium, f. temperature, g. electrical conductivity and h. total suspended solids.

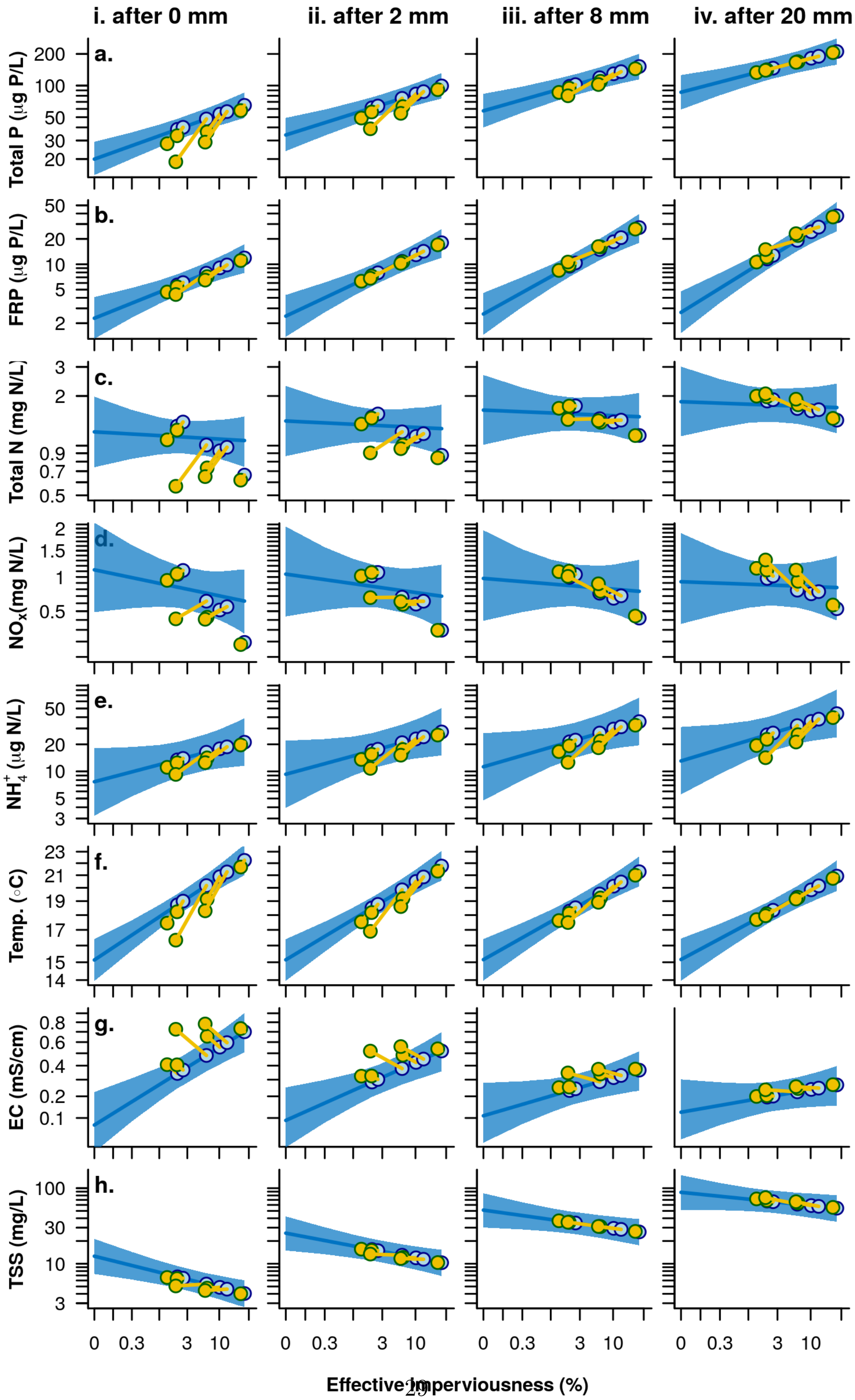

i. after 0 mm
ii. after 2 mm
iii. after 8 mm
iv. after 20 mm
a.
b.
c.
d.
e.
f.
g.
h.
Total P (µg P/L)
200
100
50
30
20
FRP (µg P/L)
50
20
10
5
2
Total N (mg N/L)
3
2
0.9
0.7
0.5
NOx(mg N/L)
2
1.5
1
0.5
NH4+ (µg N/L)
50
20
10
5
3
Temp. (○C)
23
21
19
17
15
14
EC (mS/cm)
0.8
0.6
0.4
0.2
0.1
TSS (mg/L)
100
30
10
3
0
0.3
3
10
Effective imperviousness (%)

Figure 5. Response of the eight water quality variables to effective imperviousness (EI) among the study sites (line = mean trend with EIS1, shaded polygon = 95% credible intervals) after 4 levels of rainfall in preceding 24 h (i. 0 mm, ii. 2 mm, iii. 8 mm, iv. 20 mm). For each of the 6 experimental sites, a blue point designates the concentration at the EIS1 (degrd, effective imperviousness assuming no stormwater control) of that site at the end of the study. The yellow point indicates the EIS (degrd - restr, effective imperviousness accounting for stormwater control performance) achieved at the end of the study. The blue lines (and credible intervals) therefore represent the degradation trajectories, and the yellow lines represent the restoration trajectories. For c. TN and d. NOx, the trend with EI was calculated with the septic effect set at the mean septic tank density among the study catchments, while the predictions for each experimental site were made with the septic value for that site.

4 Discussion

Like most studies of urban impacts on streams, we inferred the degrading effect of urban stormwater drainage from spatial variation among streams. In this study, we have increased confidence in this 'space-for-time' inference (sensu Pickett, 1989) by experimentally reducing drainage connection, and demonstrating changes in contaminant concentrations and temperatures, compared to control and reference sites. By quantifying both the degrading effect of stormwater runoff and the restorative effect of stormwater control in the same measure—effective imperviousness, as adapted by Walsh et al. (2022) —we have been able to directly compare the trajectories of degradation and restoration. And by casting our experiment in a Before-After-Control-Reference-Impact design, we have placed observed changes in the contexts of departure from degraded state, and approach to reference state.

The effects on stream water quality in response to SCMs that we have demonstrated provide insights into the mechanisms by which stormwater degrades stream water quality and by which SCMs mitigate that degradation, and potentially non-stormwater-related impacts as well (Table 2). Here we first discuss those mechanisms, before considering implications for better SCM design and for urban water management and protection of streams and rivers.

4.1 How stormwater degrades stream water quality, and how SCMs mitigate their effect

Stormwater drainage is as much a dry weather problem as a rain-related problem. The conventional approach to stormwater drainage manages the risk of urban flooding, but in doing so, makes the risk of downstream environmental damage a certainty. SCMs installed in this experiment had the greatest influence on stream water quality during dry weather. There were two likely causes for this effect. First, baseflows, diminished by covering of catchment soils by impervious surfaces, were likely augmented by SCMs, as evidenced by reduced temperatures and increased EC. While some of this effect may be filtered flows through the pipe network, and exfiltrated flows through gravel-filled trenches and other elements of the urban karst (Bonneau et al., 2018), increased EC suggests that exfiltrated water from SCMs also increased saline groundwater flows into the streams. These increased flows likely diluted and cooled baseflows. Second, where we were able to install SCMs that intercepted runoff from entire sub-catchments of pipes that convey runoff from roofs and roads, SCMs prevented dry-weather spills from

Table 2. A summary of degrading mechanisms relevant to our study streams, and evidence from this study for any mitigating effect of stormwater control measures. @ >p(- 2) * >p(- 2) * @ **Degrading mechanism** & **Mitigated by stormwater control measures**
STORMWATER (*degrd* effect) &

1. Dry weather spills or septic seepage to stormwater

& Yes: reduced dry-weather P, N

1. Frequent hydraulic and chemical disturbance from storm runoff

& Yes: reduced P, N, temperature after up to 8 mm rain

1. Reduced baseflows

& Yes: reduced dry-weather temperature (and increased EC)

1. Incision and widening from increased stream power

& Unclear: lack of SCM impact on TSS suggests SCMs may not have mitigated this effect: in-channel vs. catchment sources need further investigation

1. Warmer, lighter in-stream conditions (resulting from 3 and 4)

& Yes: reduced temperature

1. Loss of fine to coarse sediments (resulting from 4), perhaps leading to reduced TSS

& No: no change in TSS
SEPTIC TANK/GREY WATER SEEPAGE (septic effect) &

1. To stormwater drains

& Yes: reduced N

1. To groundwater

& Possibly: reduced N
<u>**CLIMATE CHANGE (time effect)**</u> &

1. Potential reduced dry weather flows

& Yes: reduced dry-weather temperature, increased EC

1. Warming

& Yes: reduced temperature

flowing to the stream. This was true even for the several end-of-pipe SCMs that were suboptimal in that they only had the capacity to retain and treat runoff effectively up to 1-2 mm of rain, but by providing dry-weather interception, they effectively intercepted any dry-weather spills that would otherwise have flowed to the stream.

The saline groundwater in the study catchments is likely a legacy of historic septic tank seepage. Most properties in the experimental catchments were sewered between 1985 and 1995 (South East Water, 2020; Yarra Valley Water, 2020), but some septic tanks remain throughout all catchments, as the backlog of sewer connections remains an ongoing endeavor (Victorian Auditor-General's Office, 2006). It is therefore possible that the SCM-induced increase in EC is temporary, and the legacy salts may flush out in the future.

A second legacy effect may in part explain the lack of effect of SCMs on TSS. While some of the higher TSS concentrations found in the reference streams may result from runoff from unsealed (gravel) roads, which typically drain directly to streams in some locations, it is more likely, particularly during dry weather, a result of the high bedloads of silt in these streams, which are likely mobilized and remain in suspension under relatively low flows. In contrast, the beds and banks of the control and experimental streams are dominated by consolidated clays, which are less able to be mobilized during dry weather and small rainfall events. However, during larger events, these clay channels continue to be eroded and incised, as observed during this study. Better understanding the sediment dynamics (both instream and within the catchment) of catchments subject to implementation of SCMs requires further investigation.

Stormwater drainage is known to heat stream waters (Somers et al., 2013) and this can be exacerbated by SCMs that retain shallow surface water such as wet ponds (Selbig & Bannerman, 2008). The SCMs employed in this study that released water to the stream were predominately sub-surface filtration systems without standing water, and these had a substantial cooling effect on dry weather flows in the streams (Figure 3f), cooling maximum temperatures in Ln to close to that of reference streams. Filtration-based stormwater control, thus has the potential not only to mitigate the thermal impacts of stormwater runoff (4-5 C° warming in summer), but also the impacts of climate warming, which are likely to be somewhat less (IPCC, 2021).

The SCMs employed in our study were designed to reduce total runoff volume and restore the quality and quantity of reduced baseflows. They were less successful in achieving the former than the latter (Walsh et al 2021), and yet resulted in substantial reductions in concentrations of N and P, suggesting that the observed effects were a result of treatment rather than volume reduction, as posited by Jefferson et al. (2017). We also found no evidence of retained nutrients being released during low flows, as observed in ponds and wetlands (Duan et al., 2016).

Most SCMs implemented in this study receive runoff from impervious surfaces through pipes or sealed drains, with little or no contribution from pervious runoff. Reductions in N concentrations resulting from SCMs must primarily result from either treatment of piped flows or dilution of stream flows with low-concentration filtered water. It is unlikely that the SCMs actively treated nitrate-rich groundwater emanating from septic fields in the catchments. It is possible that some high-nutrient seepage from septic fields could have leaked into stormwater drains. A more likely mechanism for treatment associated with septic tanks arises from the occurrence of septic systems in the catchment that treat only black water, with grey water drained informally to surrounding soils (Victorian Auditor-General's Office, 2006). In one property in the Ln catchment, we found that, when the catchment's roads were sealed and curbed in 2005, the grey water from such a system was diverted to the stormwater system. To deal with that problem, we constructed an infiltration system specifically for those grey water flows. It is likely that similar grey-water flows connected to the stormwater occur undetected in any of our experimental catchments, which our downslope raingardens would be intercepting.

4.2 How to better design and implement SCMs

The SCMs implemented in this project resulted in reduction of nutrient concentrations and water temperatures to levels similar to reference sites, in dominant flow conditions, in several of the experimental catchments. However, phosphorus concentrations and temperature remained substantially higher than reference concentrations during large rain events, in all except the streams of the catchments in which substantial reductions in EI_S were achieved (Figure 4). EI_S reductions achieved by the experiment were less than originally aimed for (Walsh et al., 2022), because we were unable to implement SCMs that intercepted runoff from many catchment impervious surfaces, and many SCMs had insufficient retention capacity. The experimental SCMs fell short in two ways: insufficient coverage of impervious surfaces, and insufficient retention capacity. Addressing these two shortcoming requires an ambitious and determined approach to SCM implementation.

First, SCMs need to be applied at every scale, from the individual land parcel (i.e. households) to end-of-pipe systems, designed to deal with runoff from the impervious areas in the catchment that were not able to be dealt with by at-source or intermediate systems. Strategies that rely entirely on large, centralized end-of-pipe SCMs will likely fail, as they will (i) likely be unable to ensure effective treatment with the large hydraulic loading they receive, (ii) likely be distant from demands for their water, thus reducing their effectiveness in load reduction, and (iii) potentially cause perverse effects, such as contributing to heating of waters subsequently discharged to receiving waters (Stajkowski et al., 2021). Similarly, strategies based only on at-source application of SCMs will likely fail, because inevitably space or other constraints will preclude application of SCMs to some impervious areas, as we observed in this experimental intervention (Walsh et al., 2015). A comprehensive, integrated approach is therefore needed, so that upstream systems reduce the hydraulic loading on downstream systems, and downstream systems act as insurance for upstream SCMs that fail, or for runoff from untreated impervious areas. In some instances it may be appropriate to purchase private land, where that land would allow implementation of an SCM that is critical to dealing with otherwise unmitigated runoff.

Increasing retention capacity depends on the amount of demand for harvested stormwater (Walsh et al., 2016), as well as maximizing opportunities for evapotranspiration. Maximizing the amount of demand requires end-uses which are regular (e.g. indoor uses) rather than seasonal (e.g. irrigation). Without these regular demands, rainwater and stormwater harvesting storages will remain full for much of the year (Mitchell et al., 2007), leading to frequent discharge of unmitigated runoff. SCMs should therefore be positioned to permit distribution of water from them to meet demands. This is typically done by locating rainwater tanks at the land parcel scale. Larger scale, more centralized application (i.e. central storage with reticulation back to demands) will be difficult to achieve in existing urban areas, but may be possible where redevelopment or new development is being proposed.

Evapotranspiration can be actively increased (by harvesting and then irrigation of green space), but increases can also be achieved through passively-irrigated street-tree pits (Luketich et al., 2019), or through designing stormwater treatment systems to maximize exfiltration and planting deep-rooted, high water-demand vegetation (typically trees) nearby. Western et al. (2021) demonstrated that this strategy can not only reduce the overall runoff volume, but can protect urban vegetation from water stress, with benefits for urban amenity and livability.

One important concern not considered by this study is the effectiveness of SCMs in intercepting and retaining micropollutants, microbes, heavy metals, hydrocarbons and other emerging pollutants of concerns. Many of these contaminants can be highly mobile, with several studies showing varied effectiveness of SCMs in retaining them (LeFevre et al., 2015; Schmitt et al., 2015). Widespread implementation of SCMs in a catchment should ideally follow a detailed audit of the pollutant profile, allowing the design of control measures to be optimized to treat the contaminants of concern. Attention will also need to be paid to the long-term water quality treatment performance of SCMs, including the need to maintain filter media and replace them before chemical saturation, to avoid breakthrough.

4.3 Implications for stream restoration and protection

The prediction of in-stream water quality as a function of EI_S lends strength to the proposition that ecologically successful restoration of urban streams (sensu Palmer et al., 2005) requires management intervention at a catchment scale to match the scale of the stormwater drainage impacts, rather than small-scale restoration of stream habitat (Bernhardt & Palmer, 2011). The relatively small effect on concentrations of the channel excavation upstream of Ln (i.e. the *channel* effect) compared to the effects of SCMs further points to a relatively small influence of local disturbances compared to the catchment-scale disturbance of urban stormwater drainage.

Restoration of water quality and flow regime is a fundamental basis for restoration of ecological structure and function, and retention and treatment of stormwater is critical to achieving that. Our study has shown that appropriately designed SCMs applied at a range of scales can contribute predictably to restoring water quality (and through indirect evidence, dry-weather flows) in degraded urban streams. Because EI_S integrates measures of SCM performance in its formulation (Walsh et al., 2022), it can be used to assess the effects of potential SCM designs and implementation strategies, to predict in-stream water quality responses and prioritize management actions.

For example, the combination of large-scale harvesting and infiltration systems implemented in the Ln catchment was sufficient to reduce TP concentrations during dominant flow conditions, to meet the Victorian Government environmental protection objective for TP (Figure 3A). The potential for alternative SCM design to meet such an objective can be predicted by estimation of the resulting EIS . However, the less certain achievement of the TN objective (Figure 3C) and the failure to meet the EC objective (Figure 3g) further illustrate the complexity of urban impacts. In our study catchments, which were urbanized in the 1970s, but not sewered until ~20 years later, water quality responses to mitigation of stormwater impacts are complicated by legacy effects, particularly those of septic tanks.

Variation in nitrogen concentration, in particular its dominant form, NO_x, among our study sites, including reference sites, was most strongly explained by septic tank density. Two of the reference sites had the highest septic tank density of all catchments (Table 1). Despite high septic density, and high NO_x concentrations, the reference sites Sa and Ol retained a high degree of ecological integrity: diverse invertebrate communities dominated by sensitive taxa (Walsh, 2004), leaf breakdown primarily mediated by shredder species (Imberger et al., 2008), and low-biomass algal assemblages (Taylor et al., 2004) dominated by eutrophic species (Newall & Walsh, 2005). These biotic attributes of the reference streams, together with the low concentrations of reactive phosphorus (Figure 3b), suggest that phosphorus, rather than nitrogen, limits primary productivity, with elevated nitrogen concentrations having little ecological effect. The substantial reductions in P that were achieved by SCM implementation in Ln, Ls, L4 and D8 are therefore likely to influence ecological change in these study streams (the subject of ongoing study). Although SCM implementation reduced in-stream nitrogen concentrations, further reductions, in urban and reference streams alike, will likely require retirement of septic tanks in these catchments.

The elevated EC in the urban study catchments is likely to limit the colonization and persistence of some salt-sensitive species (e.g. Kefford, 2018), and the increase caused by SCM implementation is likely to exacerbate this limitation. Spatial variation in EI (*degrd*) explained some of the variation in EC (Figure 2g, 3g), likely caused in part by elevated salt concentrations in impervious runoff from the leaching and weathering of built surfaces, particularly concrete (Kaushal et al., 2017). However, the SCM-induced, dry-weather increases in EC suggest elevated groundwater salinity, likely a legacy of historic septic tanks, but potentially a legacy of pre-urban agricultural activities as well. Infiltration of urban stormwater runoff (which is less saline, as evidenced by reduced EC with higher rainfall) into catchment soils is likely to result in long-term dilution of groundwater flows, over longer time frames than was possible in this study. SCMs are therefore most likely an appropriate tool for addressing this legacy problem for a long-term achievement of the EC target.

Such legacy effects, and indeed contemporary impacts such as inadequate sewage treatment and disposal (Piffer et al., 2021), are likely to be common in many urban catchments globally. Where problems such as inadequately treated sewage or industrial effluent are present, they should clearly be management priorities. However addressing such problems without recognition of the problem of stormwater misses an opportunity to more comprehensively restore stream ecological structure and function. Our study has demonstrated that stream water quality is impaired by small areas of conventionally drained impervious surfaces, and SCMs designed to retain and lose or treat stormwater runoff from all or near-all catchment impervious surfaces is required to mitigate such impairment to a level approaching the pre-urban state. Stormwater management to such a standard is being implemented in priority areas of the Melbourne region (Melbourne Water, 2018). While broader adoption of such an objective is likely to be challenging under dominant urban water management policy and practice, the potential co-benefits of alternative stormwater management can be large (Walsh et al., 2016).

Protection and ecologically successful restoration of river and stream ecosystems are likely to require approaches to stormwater control beyond those that currently dominate (Walsh et al., 2016). If river and stream protection is an objective, more ambitious targets are required than, for instance, those of most Mid-Atlantic municipalities, which have a goal of achieving 10–20% of the landscape drain runoff through SCMs by 2030 (Pennino et al., 2016). Stream protection will require changing standard drainage practice so that conventional drainage is no longer the default. It will require near-100% of impervious surfaces draining to SCMs (noting that SCMs should be designed for interception of runoff from impervious surfaces, not from entire landscapes: Pennino et al. (2016)). It will require SCMs in treatment trains at a range of scales from land-parcel to the largest sub-catchment that enters a stream, and with the terminal treatment supported as much as possible by upstream treatments.

Acknowledgments

The primary research activities of this project were funded by: the Australian Research Council (ARC) Linkage Program (LP0883610 and LP130100295, to CJW and TDF) with partner organizations Melbourne Water (MW) and Yarra Ranges Council (YRC); and the ARC Future Fellowship Program (FT10010044, to TDF). MW also funded research activities through the Cooperative Research Centres for Freshwater Ecology and Catchment Hydrology, and the Melbourne Waterway Research-Practice Partnership. The Smart Water Fund of the Victorian Water Trust also funded research activities.

Programs for planning, allocating, building and maintaining stormwater control measures (SCMs) in the L4 catchment were funded by: a) the Victorian Government through the Smart Water Fund of the Victorian Water Trust, the Victorian Urban Stormwater and Recycling Fund, and through Melbourne Water as part of the Port Phillip and Westernport Threatened Catchments program and Living Rivers program; b) MW; c) the Australian Government through the Caring for Our Country Investment Fund, administered by the Port Phillip and Westernport Catchment Management Authority; d) YRC; e) Yarra Valley Water. Similar programs in the D8 catchment were funded and managed by MW in collaboration with Knox City Council and South East Water.

MW is a corporation wholly owned by the Government of Victoria, with statutory responsibilities for the management of rivers and other waterways of the Melbourne region. They fund public-good research to further the environmental sustainability of their operations. The research and the University of Melbourne positions held by CJW, MJB, DGB, and SJI are all funded by the Melbourne Waterway Research-Practice Partnership, between the University of Melbourne and Melbourne Water (MW). In a previous role, TDF received payment as for his contribution to the engineering design of the Wicks Reserve Infiltration System, subsequently constructed by Knox City Council (KCC) and MW in the D8 catchment. Neither the authors' partnership with and funding by MW, nor past work with KCC, alter their adherence to AGU's ethical guidelines.

We thank Graham Rooney, Rhys Coleman, Toby Prosser and Sharyn Rossrakesh from MW for their steadfast support of this project. Belinda Hatt, Sally Boer, Monica Tewman, Ashley Macqueen, Tony Lovell, Mike Sammonds, Carleen Mitchell, Rob James, and Genevieve Hehir assisted with collection of samples. Belinda Hatt, Sandra Sdraulig, Ashley Liang and Tina Hines conducted the laboratory analyses. We thank Lee Hazel for guidance with modeling, Rhys Eddy for sewerage information and Vaughn Grey, Matt Potter, Rhys Coleman and Judy Blackbeard for comments on the manuscript.

Sampling in reference streams was conducted with Parks Victoria research permits (10007079, 10007127 and 10008347).

Open Research

The R scripts/Markdown data used for all analyses and to produce all documents of the study are available at the Open Science Framework via a link for peer review at https://osf.io/4ywvq/?view_only=c67dfccc7f1c4fffb655b1976d1678c3.

On acceptance, the repository will be made public with a DOI with a CC-By Attribution 4.0 International licence..

References

Allaire, J., Xie, Y., McPherson, J., Luraschi, J., Ushey, K., Atkins, A., et al. (2021). *rmarkdown: Dynamic documents for R*. Retrieved from https://github.com/rstudio/rmarkdown

APHA, AWWA, & WEF. (2012). *Standard Methods for the Examination of Water and Wastewater. American Public Health Association (APHA): Washington, DC, USA* (22nd ed.). Washington, DC, USA: American Public Health Association, American Water Works Association, Water Environment Federation.

Baruch, E. M., Voss, K. A., Blaszczak, J. R., Delesantro, J., Urban, D. L., & Bernhardt, E. S. (2018). Not all pavements lead to streams: variation in impervious surface connectivity affects urban stream ecosystems. *Freshwater Science*, *37*(3), 673–684. https://doi.org/10.1086/699014

Bedan, E. S., & Clausen, J. C. (2009). Stormwater runoff quality and quantity from traditional and low impact development watersheds. *JAWRA Journal of the American Water Resources Association*, *45*(4), 998–1008. https://doi.org/10.1111/j.1752-1688.2009.00342.x

Bell, C. D., McMillan, S. K., Clinton, S. M., & Jefferson, A. J. (2016). Characterizing the effects of stormwater mitigation on nutrient export and stream concentrations. *Environmental Management*, *59*(4), 604–618. https://doi.org/10.1007/s00267-016-0801-4

Bernhardt, E. S., & Palmer, M. A. (2011). River restoration: the fuzzy logic of repairing reaches to reverse catchment scale degradation. *Ecological Applications*, *21*(6), 1926–1931. https://doi.org/10.1890/10-1574.1

Bettez, N. D., & Groffman, P. M. (2012). Denitrification potential in stormwater control structures and natural riparian zones in an urban landscape. *Environmental Science and Technology*, *46*(20), 10909–10917. https://doi.org/10.1021/es301409z

Bonneau, J., Fletcher, T. D., Costelloe, J. F., Poelsma, P. J., James, R. B., & Burns, M. J. (2018). Where does infiltrated stormwater go? Interactions with vegetation and subsurface anthropogenic features. *Journal of Hydrology*, *567*, 121–132. https://doi.org/10.1016/j.jhydrol.2018.10.006

Bratieres, K., Fletcher, T. D., Deletic, A., & Zinger, Y. (2008). Nutrient and sediment removal by stormwater biofilters: a large-scale design optimisation study. *Water Research*, *42*(14), 3930–3940. https://doi.org/10.1016/j.watres.2008.06.009

Burns, M. J., Fletcher, T. D., Walsh, C. J., Ladson, A. R., & Hatt, B. E. (2012). Hydrologic shortcomings of conventional urban stormwater management and opportunities for reform. *Landscape and Urban Planning*, *105*(3), 230–240. https://doi.org/10.1016/j.landurbplan.2011.12.012

Carpenter, B., Gelman, A., Hoffman, M., Lee, D., Goodrich, B., Betancourt, M., et al. (2017). Stan: a probabilistic programming language. *Journal of Statistical Software, Articles*, *76*(1), 1–32. https://doi.org/10.18637/jss.v076.i01

Crase, B., Liedloff, A. C., & Wintle, B. A. (2012). A new method for dealing with residual spatial autocorrelation in species distribution models. *Ecography*, *35*(10), 879–888. Journal Article.

Duan, S., Newcomer-Johnson, T., Mayer, P., & Kaushal, S. (2016). Phosphorus retention in stormwater control structures across streamflow in urban and suburban watersheds. *Water*, *8*(9), 390. https://doi.org/10.3390/w8090390

EPA Victoria. (2021). *Guide to the Environmental Reference Standard. Publication 1992* (Report). Retrieved from https://www.epa.vic.gov.au/about-epa/publications/1992

Fletcher, T. D., Shuster, W., Hunt, W. F., Ashley, R., Butler, D., Arthur, S., et al. (2014). SUDS, LID, BMPs, WSUD and more—the evolution and application of terminology surrounding urban drainage. *Urban Water Journal*, *12*(7), 525–542. https://doi.org/10.1080/1573062x.2014.916314

Gelman, A., Carlin, J. B., Stern, H. S., & Rubin, D. B. (2004). *Bayesian Data Analysis* (2nd ed.). Chapman; Hall/CRC.

Gomi, T., Sidle, R. C., & Richardson, J. S. (2002). Understanding Processes and Downstream Linkages of Headwater Systems. *BioScience*, *52*(10), 905. https://doi.org/10.1641/0006-3568(2002)052%5b0905:upadlo%5d2.0.co;2

Grimm, N. B., Faeth, S. H., Golubiewski, N. E., Redman, C. L., Wu, J., Bai, X., & Briggs, J. M. (2008). Global change and the ecology of cities. *Science*, *319*(5864), 756–760. https://doi.org/10.1126/science.1150195

Guo, D., Lintern, A., Webb, J. A., Ryu, D., Bende-Michl, U., Liu, S., & Western, A. W. (2020). A data-based predictive model for spatiotemporal variability in stream water quality. *Hydrology and Earth System Sciences*, *24*(2), 827–847. https://doi.org/10.5194/hess-24-827-2020

Hale, R. L., Turnbull, L., Earl, S. R., Childers, D. L., & Grimm, N. B. (2014). Stormwater infrastructure controls runoff and dissolved material export from arid urban watersheds. *Ecosystems*, *18*(1), 62–75. https://doi.org/10.1007/s10021-014-9812-2

Hatt, B. E., Fletcher, T. D., Walsh, C. J., & Taylor, S. L. (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. https://doi.org/10.1007/s00267-004-0221-8

Hill, P., Mein, R., & Siriwardena, L. (1998). *How much rainfall becomes runoff? Loss modelling for flood estimation* (Report No. Report 98/5). Cooperative Research Centre for Catchment Hydrology. Retrieved from https://www.ewater.org.au/archive/crcch/archive/pubs/pdfs/industry199805.pdf

Hopkins, K. G., Loperfido, J. V., Craig, L. S., Noe, G. B., & Hogan, D. M. (2017). Comparison of sediment and nutrient export and runoff characteristics from watersheds with centralized versus distributed stormwater management. *Journal of Environmental Management*, *203*, 286–298. https://doi.org/10.1016/j.jenvman.2017.07.067

Imberger, S. J., Walsh, C. J., & Grace, M. R. (2008). More microbial activity, not abrasive flow or shredder abundance, accelerates breakdown of labile leaf litter in urban streams. *Journal of the North American Benthological Society*, *27*(3), 549–561. Journal Article. https://doi.org/10.1899/07–123.1

IPCC. (2021). *Climate change 2021: The physical science basis. Contribution of working group i to the sixth assessment report of the intergovernmental panel on climate change*. book, Cambridge, UK: Cambridge University Press.

Jefferson, A. J., Bhaskar, A. S., Hopkins, K. G., Fanelli, R., Avellaneda, P. M., & McMillan, S. K. (2017). Stormwater management network effectiveness and implications for urban watershed function: A critical review. *Hydrological Processes*, *31*(23), 4056–4080. https://doi.org/10.1002/hyp.11347

Kaushal, S. S., Duan, S., Doody, T. R., Haq, S., Smith, R. M., Newcomer Johnson, T. A., et al. (2017). Human-accelerated weathering increases salinization, major ions, and alkalinization in fresh water across land use. *Applied Geochemistry*, *83*, 121–135. https://doi.org/10.1016/j.apgeochem.2017.02.006

Kaushal, S. S., Wood, K. L., Galella, J. G., Gion, A. M., Haq, S., Goodling, P. J., et al. (2020). Making 'chemical cocktails'—evolution of urban geochemical processes across the periodic table of elements. *Applied Geochemistry*, *119*, 104632. https://doi.org/10.1016/j.apgeochem.2020.104632

Kefford, B. J. (2018). Why are mayflies (Ephemeroptera) lost following small increases in salinity? Three conceptual osmophysiological hypotheses. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *374*(1764), 20180021. https://doi.org/10.1098/rstb.2018.0021

Kim, D., & Song, S.-K. (2019). The Multifunctional Benefits of Green Infrastructure in Community Development: An Analytical Review Based on 447 Cases. *Sustainability*, *11*(14), 3917. https://doi.org/10.3390/su11143917

Lamond, J., & Everett, G. (2019). Sustainable Blue-Green Infrastructure: A social practice approach to understanding community preferences and stewardship. *Landscape and Urban Planning*, *191*, 103639. https://doi.org/10.1016/j.landurbplan.2019.103639

LeFevre, G. H., Paus, K. H., Natarajan, P., Gulliver, J. S., Novak, P. J., & Hozalski, R. M. (2015). Review of Dissolved Pollutants in Urban Storm Water and Their Removal and Fate in Bioretention Cells. *Journal of Environmental Engineering*, *141*(1), 04014050. https://doi.org/10.1061/(asce)ee.1943-7870.0000876

Luketich, A. M., Papuga, S. A., & Crimmins, M. A. (2019). Ecohydrology of urban trees under passive and active irrigation in a semiarid city. *PLOS ONE*, *14*(11), e0224804. https://doi.org/10.1371/journal.pone.0224804

Melbourne Water. (2018). *Healthy waterways strategy 2018* (Report). Melbourne Water. Retrieved from https://www.melbournewater.com.au/about/strategies-and-reports/healthy-waterways-strategy

Mitchell, V. G., Deletic, A., Fletcher, T. D., Hatt, B. E., & McCarthy, D. T. (2007). Achieving multiple benefits from stormwater harvesting. *Water Science and Technology*, *55*(4), 135–144. https://doi.org/10.2166/wst.2007.103

Newall, P., & Walsh, C. J. (2005). Response of epilithic diatom assemblages to urbanization influences. *Hydrobiologia*, *532*(1-3), 53–67. https://doi.org/10.1007/s10750-004-9014-6

Palmer, M. A., Bernhardt, E. S., Allan, J. D., Lake, P. S., Alexander, G., Brooks, S., et al. (2005). Standards for ecologically successful river restoration. *Journal of Applied Ecology*, *42*(2), 208–217. https://doi.org/10.1111/j.1365-2664.2005.01004.x

Pennino, M. J., McDonald, R. I., & Jaffe, P. R. (2016). Watershed-scale impacts of stormwater green infrastructure on hydrology, nutrient fluxes, and combined sewer overflows in the mid-Atlantic region. *Science of The Total Environment*, *565*, 1044–1053. https://doi.org/10.1016/j.scitotenv.2016.05.101

Petrucci, G., Rioust, E., Deroubaix, J.-F., & Tassin, B. (2013). Do stormwater source control policies deliver the right hydrologic outcomes? *Journal of Hydrology*, *485*, 188–200. https://doi.org/10.1016/j.jhydrol.2012.06.018

Pickett, S. T. A. (1989). Space-for-time substitution as an alternative to long-term studies. In G. E. Likens (Ed.), *Long-term studies in ecology: Approaches and alternatives* (pp. 110–135). Springer, New York. https://doi.org/10.1007/978-1-4615-7358-6_5

Piffer, P. R., Tambosi, L. R., Ferraz, S. F. B., Metzger, J. P., & Uriarte, M. (2021). Native forest cover safeguards stream water quality under a changing climate. *Ecological Applications*. https://doi.org/10.1002/eap.2414

R Core Team. (2020). *R: A language and environment for statistical computing*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/

Roy, A. H., Rhea, L. K., Mayer, A. L., Shuster, W. D., Beaulieu, J. J., Hopton, M. E., et al. (2014). How much is enough? Minimal responses of water quality and stream biota to partial retrofit stormwater management in a suburban neighborhood. *PLoS ONE*, *9*(1), e85011. https://doi.org/10.1371/journal.pone.0085011

Schmitt, N., Wanko, A., Laurent, J., Bois, P., Molle, P., & Mosé, R. (2015). Constructed wetlands treating stormwater from separate sewer networks in a residential Strasbourg urban catchment area: Micropollutant removal and fate. *Journal of Environmental Chemical Engineering*, *3*(4), 2816–2824. https://doi.org/10.1016/j.jece.2015.10.008

Selbig, W. R., & Bannerman, R. T. (2008). *A comparison of runoff quantity and quality from two small basins undergoing implementation of conventional and low-impact-development (LID) strategies: Cross Plains, Wisconsin, water years 1999–2005* (U.S. Geological Survey Scientific Investigations Report 2008–5008). Retrieved from https://pubs.usgs.gov/sir/2008/5008/pdf/sir_2008-5008.pdf

Somers, K. A., Bernhardt, E. S., Grace, J. B., Hassett, B. A., Sudduth, E. B., Wang, S., & Urban, D. L. (2013). Streams in the urban heat island: Spatial and temporal variability in temperature. *Freshwater Science*, *32*(1), 309--326. Journal Article. https://doi.org/10.1899/12-046.1

South East Water. (2020). *South East Water asset web map* (Report). South East Water. Retrieved from https://webmap.southeastwater.com.au/sewgis/

Stajkowski, S., Laleva, A., Farghaly, H., Bonakdari, H., & Gharabaghi, B. (2021). Modelling dry-weather temperature profiles in urban stormwater management ponds. *Journal of Hydrology*, *598*, 126206. https://doi.org/10.1016/j.jhydrol.2021.126206

Stan Development Team. (2020a). RStan: The R interface to Stan. Retrieved from http://mc-stan.org/

Stan Development Team. (2020b). Stan Modeling Language Users Guide and Reference Manual. Retrieved from http://mc-stan.org

Taylor, S. L., Roberts, S. C., Walsh, C. J., & Hatt, B. E. (2004). Catchment urbanisation and increased benthic algal biomass in streams: linking mechanisms to management. *Freshwater Biology*, *49*(6), 835–851. https://doi.org/10.1111/j.1365-2427.2004.01225.x

Vehtari, A., Gelman, A., & Gabry, J. (2017). Practical bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, *27*(5), 1413–1432.

Victorian Auditor-General's Office. (2006). *Protecting our environment and community from failing septic tanks*. book, Melbourne: Victorian Government Printer. Retrieved from https://www.parliament.vic.gov.au/papers/govpub/VPARL2003-06No205.pdf

Vörösmarty, C. J., McIntyre, P. B., Gessner, M. O., Dudgeon, D., Prusevich, A., Green, P., et al. (2010). Global threats to human water security and river biodiversity. *Nature*, *467*(7315), 555–561. https://doi.org/10.1038/nature09440

Walsh, C. J. (2004). Protection of in-stream biota from urban impacts: minimise catchment imperviousness or improve drainage design? *Marine and Freshwater Research*, *55*(3), 317. https://doi.org/10.1071/mf03206

Walsh, C. J., & Kunapo, J. (2009). The importance of upland flow paths in determining urban effects on stream ecosystems. *Journal of the North American Benthological Society*, *28*(4), 977–990. https://doi.org/10.1899/08-161.1

Walsh, C. J., Fletcher, T. D., & Ladson, A. R. (2005). Stream restoration in urban catchments through redesigning stormwater systems: looking to the catchment to save the stream. *Journal of the North American Benthological Society*, *24*(3), 690–705. https://doi.org/10.1899/04-020.1

Walsh, C. J., Roy, A. H., Feminella, J. W., Cottingham, P. D., Groffman, P. M., & Morgan, R. P. (2005). The urban stream syndrome: current knowledge and the search for a cure. *Journal of the North American Benthological Society*, *24*(3), 706–723. Journal Article. https://doi.org/10.1899/04-028.1

Walsh, C. J., Fletcher, T. D., & Burns, M. M. J. (2012). Urban stormwater runoff: a new class of environmental flow problem. *PLoS ONE*, *7*(9), e45814. https://doi.org/10.1371/journal.pone.0045814

Walsh, C. J., Fletcher, T. D., Bos, D. G., & Imberger, S. J. (2015). Restoring a stream through retention of urban stormwater runoff: a catchment-scale experiment in a social-ecological system. *Freshwater Science*, *34*(3), 1161–1168. https://doi.org/10.1086/682422

Walsh, C. J., Booth, D. B., Burns, M. J., Fletcher, T. D., Hale, R. L., Hoang, L. N., et al. (2016). Principles for urban stormwater management to protect stream ecosystems. *Freshwater Science*, *35*(1), 398–411. https://doi.org/10.1086/685284

Walsh, C. J., Burns, M. J., Fletcher, T. D., Bos, D. G., Kunapo, J., Poelsma, P.,

& Imberger, S. J. (2021). Linking stormwater control performance to stream ecosystem outcomes: Incorporating a performance metric into effective imperviousness/data and code. *Open Science Framework*. https://doi.org/10.17605/OSF.IO/57AZQ

Walsh, C. J., Burns, M. J., Fletcher, T. D., Bos, D. G., Kunapo, J., Poelsma, P., & Imberger, S. J. (2022). Linking stormwater control performance to stream ecosystem outcomes: Incorporating a performance metric into effective imperviousness. *PLOS Water*, in press. Journal Article. Retrieved from https://osf.io/57azq

Western, A. W., Arora, M., Burns, M. J., Bonneau, J., Thom, J. K., Yong, C. F., et al. (2021). Impacts of stormwater infiltration on downslope soil moisture and tree water use. *Environmental Research Letters*, *16*(10), 104014. https://doi.org/10.1088/1748-9326/ac1c2a

Wilson, C. E., Hunt, W. F., Winston, R. J., & Smith, P. (2015). Comparison of runoff quality and quantity from a commercial low impact and conventional development in Raleigh, North Carolina. *Journal of Environmental Engineering*, *141*(2), 05014005. https://doi.org/10.1061/(asce)ee.1943-7870.0000842

Yarra Valley Water. (2020). *Yarra Valley Water asset map* (Report). Yarra Valley Water. Retrieved from https://webmap.yvw.com.au/yvw_ext/