

RESEARCH ARTICLE

Linking stormwater control performance to stream ecosystem outcomes: Incorporating a performance metric into effective imperviousness

Christopher J. Walsh ^{*}, Matthew J. Burns, Tim D. Fletcher, Darren G. Bos, Peter Poelsma , Joshphar Kunapo , Moss Imberger 

Waterway Ecosystem Research Group, School of Ecosystem and Forest Sciences, The University of Melbourne, Burnley, Victoria, Australia

* cwalsh@unimelb.edu.au



OPEN ACCESS

Citation: Walsh CJ, Burns MJ, Fletcher TD, Bos DG, Poelsma P, Kunapo J, et al. (2022) Linking stormwater control performance to stream ecosystem outcomes: Incorporating a performance metric into effective imperviousness. *PLOS Water* 1(2): e0000004. <https://doi.org/10.1371/journal.pwat.0000004>

Editor: Soni M. Pradhanang, University of Rhode Island, UNITED STATES

Received: July 13, 2021

Accepted: November 3, 2021

Published: February 15, 2022

Peer Review History: PLOS recognizes the benefits of transparency in the peer review process; therefore, we enable the publication of all of the content of peer review and author responses alongside final, published articles. The editorial history of this article is available here: <https://doi.org/10.1371/journal.pwat.0000004>

Copyright: © 2022 Walsh et al. This is an open access article distributed under the terms of the [Creative Commons Attribution License](https://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Data Availability Statement: All data and code are available from the Open Science Framework: <https://doi.org/10.17605/OSF.IO/57AZQ>.

Abstract

Stormwater control measures, such as raingardens, tanks, or wetlands, are often employed to mitigate the deleterious effects of urban stormwater drainage on stream ecosystems. However, performance metrics for control measures, most commonly pollutant-load reduction, have not permitted prediction of how they will change stream ecosystems downstream. Stream ecosystem responses have more commonly been predicted by catchment-scale measures such as effective imperviousness (percentage of catchment with impervious cover draining to sealed drains). We adapt effective imperviousness, weighting it by a performance metric for stormwater control measures aimed at stream protection, the stream stormwater impact metric. Weighted effective imperviousness can serve as a predictor of stream response to stormwater control. We demonstrate its application in a before-after-control-reference-impact experiment aiming to test if stream health is improved by dispersed stormwater control measures. Trends in weighted effective imperviousness showed wide variation in degree of stormwater control achieved in the six experimental sub-catchments, despite similar effort in implementing control measures across the sub-catchments. Greater reductions in weighted effective imperviousness (on a log-scale, on which stream response is predicted) per unit effort were observed in smaller catchments with lower starting effective imperviousness. While implementation of control measures was sufficient to expect a stream response in at least two of the experimental sub-catchments, we did not achieve the reduction in effective imperviousness that we were aiming for. Primary limitations to success were the lack of available space in these established suburbs, particularly for final control measures near pipe outlets into streams, and a lack of demand for harvested stormwater. The use of the continuous variable, weighted effective imperviousness, to measure impact on streams, and the protracted period of SCM implementation that varied among catchments, required a new approach to modelling “before-after-control-impact” experiments, which has potentially broader application.

Funding: The research activities reported in this paper span 19 years, and were supported by many funding sources. The primary research activities of this project were funded by: the Australian Research Council (ARC, <https://www.arc.gov.au/>) Linkage Program through grants awarded to CJW and TDF (LP0883610 and LP130100295) with partner organizations Melbourne Water (MW, <https://www.melbournewater.com.au/>) and Yarra Ranges Council (YRC, <https://www.yarraranges.vic.gov.au/>); and the ARC Future Fellowship Program through a grant awarded to TDF (FT10010044). Melbourne Water additionally funded research activities, at first through the Cooperative Research Centres for Freshwater Ecology and Catchment Hydrology, and later through the Melbourne Waterway Research-Practice Partnership (<https://mwrpp.org/>). The Smart Water Fund of the Victorian Water Trust (SWF) also funded research activities. Programs for planning, allocating, building and maintaining stormwater control measures (SCMs) in the L4 catchment were funded by: the Victorian Government through the SWF, 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; Melbourne Water; the Australian Government through the Caring for Our Country Investment Fund, administered by the Port Phillip and Westernport Catchment Management Authority; Yarra Ranges Council; Yarra Valley Water (YVW, <https://www.yvw.com.au/>), who administered SCM incentive payments. Similar programs of SCM implementation in the D8 catchment were funded and managed by Melbourne Water in collaboration with Knox City Council (KCC) and South East Water. The authors collected performance data for D8 SCMs, and provided advice on some aspects of SCM design, but were not involved in program administration or implementation and received no project funding for works in the D8 catchment. None of the ARC, the Victorian or Australian Government funding bodies, or YVW played a role in the research. The authors worked with MW, YRC and KCC on the SCM implementation programs.

Competing interests: I have read the journal's policy and the authors of this manuscript have the following competing interests. The research and the University of Melbourne positions held by CJW, MJB, DGB, PP, JK, and SJI are all funded by the Melbourne Waterway Research-Practice Partnership, between the University of Melbourne and Melbourne Water (MW). MW is a corporation wholly owned by the Government of Victoria, with statutory responsibilities for the management of

Introduction

Conventional urban stormwater drainage, delivering runoff from impervious surfaces to streams and rivers through hydraulically efficient pipes and sealed drains, severely degrades receiving stream ecosystems. Globally, urban streams are characterized by loss of sensitive species and shifts in ecological function [1], which result from alterations to flow regimes, channel form [2, 3], and water quality [4, 5]. These in-stream changes arise because conventional stormwater drainage increases the frequency and magnitude of disturbance. The disturbances experienced by channels and their in-stream biota are both hydraulic, arising from larger, more frequent high-flow events, and chemical, arising from the complex cocktails of pollutants associated with impervious runoff [6]. Hydraulic changes also simplify in-stream habitat, by increasing the capacity of streams to transport sediments [7], causing the loss of habitat features such as bars, benches, and woody debris [2].

In part as a response to increased awareness of the deleterious effects of conventional stormwater drainage to receiving aquatic ecosystems, alternative drainage approaches have been advocated across many jurisdictions over the last three decades [8]. While the protection of receiving waters has been a primary objective behind such approaches from the beginning [e.g. 9, 10], the connection between performance objectives of stormwater control measures (SCMs) and receiving water response has largely been opaque. Objectives have most commonly been set as performance metrics of (annual) pollutant load reduction [from the loads produced by conventional drainage: e.g. 11–13]. Such metrics fail to address the complexity, or match the time scales, of the hydrologic and geomorphic stressors to stream ecosystems created by urban stormwater runoff [14]. However, even SCM performance metrics that are explicitly linked to mechanisms of stream degradation [e.g. 15] cannot be easily used to predict the effects of the SCMs on their receiving stream ecosystem. Such prediction requires integration of SCM performance metrics with catchment-scale predictors of stream response.

Studies of stream ecosystem response to catchment urbanization have shown a predictable decline in ecological indicators with increased catchment urban density [16]. Progress in understanding of mechanisms for this decline has been hampered by a lack of consistency in catchment metrics used as predictor variables [17]. Total catchment imperviousness (*TI*, proportion of catchment covered by impervious surfaces) has most commonly been used, but several studies have shown that stream degradation is better explained when drainage connection is accounted for in addition to impervious coverage [16, 18–20]. If we are to understand differences or similarities in response to degradation of urban streams across regions, let alone the potential for reversing that degradation, we need consistent metrics of elements of urban infrastructure that are likely to drive stream ecosystem change. The most important such element is likely the stormwater drainage system [17].

Effective imperviousness (*EI*) is the proportion of catchment area covered by effective impervious surfaces, where 'effective' connotes those surfaces that have an effect on the receiving stream. This measure arose from Leopold's [21] demonstration that hydrologic changes in urban catchments are a function of both imperviousness and the extent of sealed stormwater drainage. In its simplest formulation, only impervious surfaces with direct connection to stormwater drainage are considered effective, while those that drain informally to pervious land are considered to have no effect. Such a formulation of *EI* was a better predictor of ecological response than *TI* in streams of peri-urban eastern Melbourne, Australia, where the extent and nature of stormwater infrastructure varies within and between catchments, ranging from informal drainage to surrounding pervious land, to formal pipe networks receiving runoff from curbed roads and private properties [16]. That *EI* was a better predictor than *TI* in such a setting provided evidence that conventional drainage of impervious areas has a greater

rivers and other waterways of the Melbourne region. They fund public-good research to further the environmental sustainability of their operations. 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 PLOS Water policies. The authors have no other competing interests.

impact on stream ecosystems than informal drainage. Impervious surfaces connected to the drainage system are undoubtedly important contributors to stream degradation given the substantial changes to catchment hydrologic processes that they cause [17].

However, it does not follow from *EI* being a better predictor that informally drained impervious surfaces have no effect. While the original formulation of *EI* made this assumption, a more parsimonious interpretation of *EI* being a better predictor of stream degradation than *TI* [16], would be that the effect of informally drained surfaces is substantially smaller than that of conventionally drained surfaces. While small, the effect of an informally drained surface is likely to depend on the length and nature (e.g. roughness and permeability) of the flow path between it and the stream. Walsh and Kunapo [18] attempted to account for such variability by weighting impervious surface areas by their distance to the nearest downstream stormwater pipe, drain or stream. However, their formulation conflated the probability that an impervious surface in an area with formal stormwater drainage systems was connected to a stormwater drain, with the variable effect of informally drained surfaces. These two effects could be separated by first identifying those impervious surfaces that are directly connected to the stormwater drainage network, and then modeling the degree of stormwater retention between informally drained surfaces and the stream, and weighting their area accordingly (between 1, if directly connected to conventional drainage, and 0, if draining to surrounding impervious land without significantly altering hydrologic processes).

A similar logic could be applied to the effects of SCMs. The performance of an SCM could be used to determine an appropriate weighting of the impervious area upstream of it. Water fluxes can be modeled through SCMs of known specification, allowing calculation of *EI* like that proposed by Walsh, Fletcher [22]. They proposed weighting impervious areas that drain to SCMs by a metric of overflow frequency, but it is also possible to weight by other aspects of SCM performance, such as restoration of filtered base flow or reduction in total flow volume [15, 23]. In this paper we advance the SCM performance metric proposed by Fletcher, Walsh [23] and Walsh, Fletcher [15] by using it and its subindices to formulate several alternative variants of *EI*. We advance the proposed approach of Walsh, Fletcher [22], by devising and demonstrating a method for combining performance metrics of complex arrays of SCMs to derive the catchment *EI* variants. In considering the effects of SCMs on *EI*, we advance the approach of Roy, Rhea [24], who removed impervious area draining to SCMs from *EI*, regardless of SCM performance.

We use data from a 19-year experiment of urban stormwater impacts in small-stream catchments of eastern Melbourne, Australia. The primary aim of the experiment was to test if intensive application of dispersed SCMs can restore stream hydrology, water quality and ultimately ecological state in receiving streams [15]. The restoration experiment was predicated on the hypothesis that impervious runoff delivered through conventional drainage systems was the primary degrading process requiring remediation to permit restoration of the receiving stream ecosystem. Our study catchments differed in their starting degrees of urban density and stormwater drainage infrastructure and, over 4 years, we applied different intensities, extents, and types of stormwater control measures in two experimental catchments. We are thus able to compare variation of *EI* variants as catchment-scale indicators of SCM performance over time and between catchments.

The study was conceived as a Before-After-Control-Reference-Impact (BACRI) experiment [15]. The classical approach to Before-After-Control-Impact experiments (and their extension to include reference sites in addition to control and impact sites) has been to divide samples into before and after periods to form a binary before-after effect, and to classify reference, control and experimental/impact sites categorically [25]: the presence of an experimental impact or intervention is tested by the interaction between the before-after effect and the

control-impact (-reference) effect. However, such a categorical approach to testing the effect of SCMs in this experiment is not optimal because of the cumulative nature of SCM implementation; the continuous nature of *EI*, which was used to define reference and control catchments; and the variable degree of intervention. The protracted period of implementing the many dispersed SCMs in our study resulted from the logistical complexities of engagement with communities and authorities, and designers, builders and maintainers of the SCMs [15, 26–28]. Long periods of experimental manipulation are likely to be typical of similar endeavours [e.g. 24]. An alternative, but analogous, analytical approach is required to assess the effects of such continuous, variable interventions.

In this paper, we propose such a modelling approach and use the experiment's implementation data and two contrasting simulated response variables to test and validate the model, and to also assess if the experimental manipulation (SCM implementation) was of a sufficient magnitude to expect a detectable response in the receiving streams. The primary focus of the paper is reporting and interpreting the implementation data of the experiment, and to propose and demonstrate new methods for assessing a) the degree of dispersed catchment stormwater runoff achieved by the SCMs implemented in the experiment, and b) their effect on downstream waters. Further investigations of stream response will use these data and approaches to assess the effects of the experiment on hydrologic, water quality and ecological variables measured in the receiving streams over the period of study. Before reporting and interpreting the SCM implementation of the experiment, we propose in the following section, a new variant of effective imperviousness, weighted by a performance metric for stormwater control measures. This measure will not only aid interpretation of the experiment, but more broadly allow prediction of how streams respond to stormwater control across catchments.

A new variant of *EI*

***EI* as a catchment-scale predictor of stream ecosystem degradation.** Walsh, Fletcher [16] portrayed the degradation in streams of eastern Melbourne as a function of *EI*, predicted by a broken-stick model, which was used to estimate a minimum threshold of *EI* at which degradation was maximal and beyond which further degradation was undetectable (Fig 1). The non-linear degradation observed in response to *EI* is also well modeled by exponential decay, or by a linear model when *EI* is logarithmically scaled (Fig 1; arguably a logistic curve would be a more appropriate transformation for the proportional measure, *EI*, but as the maximum *EI* among our catchments was 25%, the simpler log transformation is a good approximation). Therefore, to assess stream responses to urban development (degradation) and SCM implementation (restoration), we propose a model of linear response to $\log_{10}(EI + 0.1)$ (with *EI* as a percentage and the addition 0.1 to permit zero values: Fig 1). The in-stream response variables, modelled by Walsh, Fletcher [16] and depicted in Fig 1, are important because they correspond to metrics used in legislated objectives for stream ecosystem protection: median filterable reactive phosphorus (Fig 1A) is comparable to 75th percentile total Phosphorus (TP) used as a water quality objective by Victoria's State Environment Protection Policy and SIGNAL score is comparable to SIGNAL2 used for biological objectives [29].

Thus *EI* can predict attainment of legislated environmental objectives following development with conventional stormwater drainage, and if it can be adapted to incorporate SCM performance, also has the potential to predict attainment of objectives as a result of stormwater management efforts to mitigate impacts.

A stream protection metric of SCM performance. The stream stormwater impact metric, *S*, estimates the extent to which an SCM mimics the catchment hydrology and water quality of appropriate reference streams that support good ecological condition. It is the mean of

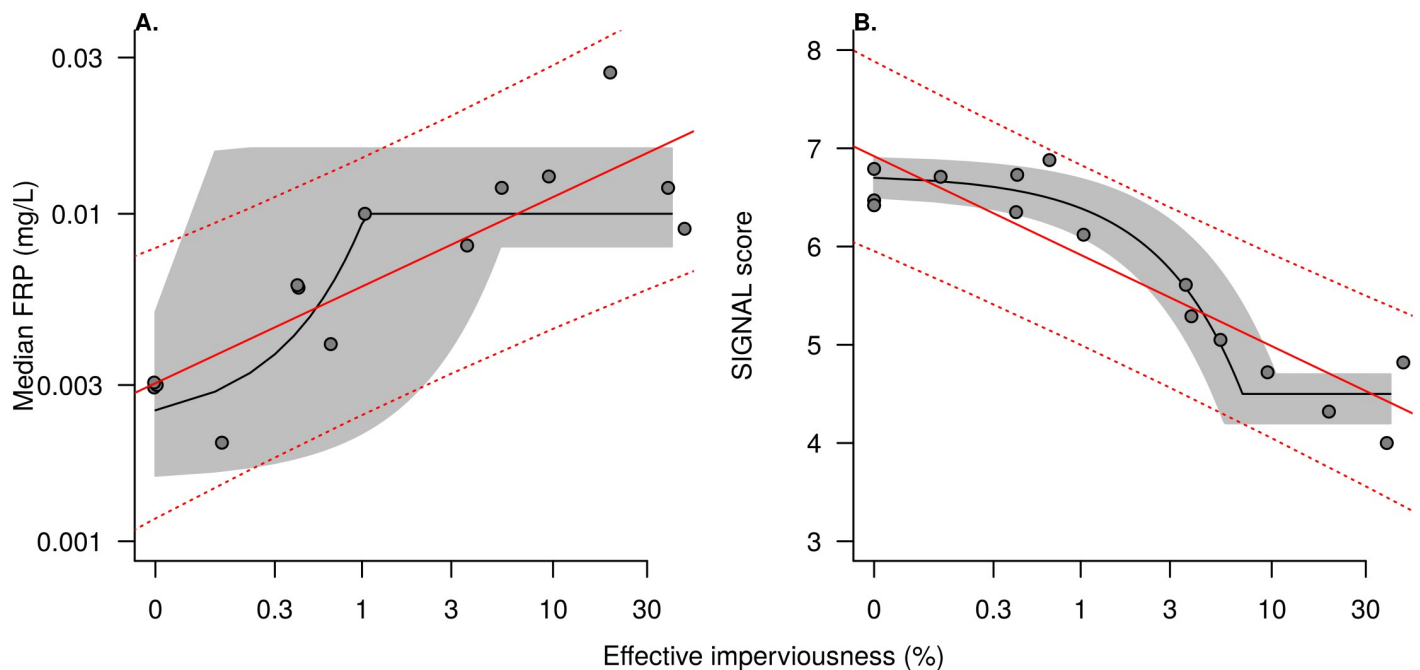


Fig 1. Examples of stream ecosystem response variables as a function of effective imperviousness (EI). (A) Median filterable reactive phosphorus (FRP) concentrations and (B) SIGNAL score (a biotic index). Piecewise regressions [solid black line with 95% confidence intervals in the grey polygon as derived by 16] are compared to linear regressions against $\log_{10}(\%EI + 0.1)$, solid red line with dotted 95% confidence intervals. Five overlapping points in (A) are shown by slight jittering.

<https://doi.org/10.1371/journal.pwat.0000004.g001>

four sub-metrics of different aspects of hydrology that can be addressed through SCM design and that have been posited as the primary drivers of stream degradation by urban stormwater runoff [14, 15, 23]. To calculate each sub-metric, a hydrologic index (I) for the impervious area draining to the SCM is modelled (using an average year of rainfall data, in this case 1965–1966) in three potential states:

- u : the impervious area conventionally drained with no SCM;
- m : the impervious area draining to the SCM, modelled based on its specifications;
- n : the area covered by impervious surfaces in its pre-urban state, inferred, in our study, from the runoff behaviour of nearby vegetated reference catchments.

Within bounds described in S1 Text, each sub-metric is the ratio of the divergence of I from the pre-urban state with the SCM in place and the divergence of I from the pre-urban state if the SCM was absent, thus scaling the effect of the SCM between worst (no stormwater control) and best (target or reference condition) case. Thus:

$$S_I = \frac{I_m - I_n}{I_u - I_n} \quad (1)$$

where S_I is the sub-metric of S based on the hydrologic index I , and the three values of I denoted by subscripts for three potential states as described above. The derivation and calculation of the four sub-metrics is described in detail in S1 Text, but briefly:

- S_R , the runoff-frequency sub-metric, assesses how closely the frequency of overflows from the SCM approaches the frequency of storm flow from the SCM's catchment before it was built over (i.e. in its pre-urban, vegetated state);

- S_F , the filtered-flow-volume sub-metric, assesses how closely the volume of water treated by the SCM and allowed to drain to the stream at an appropriate rate, matches stream baseflow generated by the SCM's catchment in its pre-urban, vegetated state;
- S_W , the water-quality sub-metric, assesses the degree to which concentrations of important pollutants (TP; total nitrogen, TN; and total suspended solids, TSS) of outflows from the SCM approach reference condition concentrations;
- S_V , the volume-reduction sub-metric, assesses the degree to which the total volume of water reaching the stream from the SCM approaches the volume of streamflow generated by the SCM's catchment before it was built over (i.e. in its pre-urban, vegetated state).

EI_S as a catchment-scale predictor of mitigation of stormwater impacts by SCMs. As S is a fraction ranging between 0 (perfectly mimicking aspects of the reference catchment flow regime) and 1 (flow regime from conventional stormwater drainage with no SCMs), it can be used to weight impervious areas draining to SCMs, leading to a variant of EI that accounts for mitigation of urban stormwater impacts by SCMs. Thus:

$$EI_S = \frac{\sum_i^n (IA_i * S_i)}{CA} \quad (2)$$

Where IA_i is the area of the i th of n groups of impervious surfaces in the catchment, weighted by S_i , and CA is the total area of the catchment. For each impervious surface not draining to an SCM, $S = 1$, and for each SCM ($S \leq 1$), all impervious surfaces that drain to it are summed to calculate IA .

We argue that our calculation of S , by averaging performance metrics that assess mitigation of putative primary drivers of stream degradation by urban stormwater runoff, provides an appropriate weighting of EI to predict stream ecosystem response to SCMs. However, it is only one of many possible ways to weight SCM performance, and alternative weightings can be heuristic for understanding any shortcomings in SCM performance, if not necessarily predictive of stream response.

For instance, EI_{S1} , assuming all SCMs have no effect ($S = 1$), and EI_{S0} , assuming all SCMs perform perfectly [i.e. $S = 0$, as used by 24], provide bounds against which to assess the overall performance of SCMs in the catchment as estimated by EI_S . If EI_{S0} and EI_S are similar (and non-zero), it shows that most of the catchment's impervious surfaces remain untreated by SCMs. If they are strongly different, and EI_S is similar to EI_{S0} then most of the SCMs mimic reference flow and water quality well, and EI weighted by each of the four sub-metrics (EI_{SR} , EI_{SF} , EI_{SW} , EI_{SV}) indicates which design objectives have contributed most strongly to the performance of EI_S .

We thus aim to use these catchment-scale measures of SCM performance to report and interpret the BACRI experiment on the effects of dispersed SCMs. We first describe the experimental design, and methods used to derive estimates of impervious coverage and SCM performance that are required to calculate the EI variants. We use the EI variants to assess the degree to which stormwater was controlled by the SCMs installed during the experiment. Finally, we propose and test a statistical model for assessing our BACRI experiment using the calculated EI variants. The proposed model has potentially broader application to other studies with continuous measures of the "control-reference-impact" effect and a protracted period of impact (in this case, SCM implementation) that varies between sites.

Methods and materials

Study catchments and experimental design

The BACRI experimental design described by Walsh, Fletcher [15] comprised seven independent catchments of different urban density and drainage connection, located in the Dandenong Ranges on the eastern fringe of the Melbourne metropolitan area (Fig 2). The original design comprised:

- three reference, forested catchments (Sa, Ly, and Ol, Fig 2), with little or no stormwater drainage infrastructure;
- three control urban catchments (Br, Fe, and D4), with streams degraded by urban stormwater drainage; and
- one experimental urban catchment (Little Stringybark Creek, L4, Fig 2), in which stormwater control measures were implemented progressively from 2009.

In 2012 and 2013, Melbourne Water initiated a parallel program of SCM implementation in D8, a larger catchment that includes D4 (Dobsons Creek, Fig 2). This new program converted D4 from a control to an experimental catchment. From 2010, the hydrologic and ecological monitoring that had been conducted at the original sites since 2001, were expanded to include D8, and three sites with independent sub-catchments upstream of L4 (Fig 2). The resulting 6 experimental catchments (the original L4, plus D4, D8, and the three tributaries upstream of L4: Ln, Ls, and L1) had different degrees of *EI* and of SCM implementation, permitting several continuous assessments of intervention effects that were not possible with the original experimental design, which had just a single experimental site.

This work was completed with the approval of the University of Melbourne's Human Research Ethics Committee (HREC No: 0720064), with Parks Victoria research permits (10007079, 10007127 and 10008347).

Estimation of effective imperviousness variants

Impervious coverage and drainage connection. Methods for impervious coverage estimation are described in detail in S2 Text. Briefly, impervious coverage in the study catchments was first estimated from 2009 4-band aerial photos and LiDAR data using a semi-automated method. These impervious coverage estimates were corrected and refined for our study catchments, by ground-truthing, and visual inspection of aerial images from 2009 to 2019 (www.nearmap.com). This imagery, together with the 2000 and 2004 imagery used in earlier studies [16, 18], was used to estimate dates of constructions or demolitions of impervious surfaces in the study catchments (there was no significant urban expansion in any of the catchments over the study period, but there was some infill urban development). These data were used to map and date all impervious surfaces in the study catchments from 2000 to 2019.

Connection of impervious surfaces to stormwater drainage was determined for all but two catchments, over the 19 years of monitoring and working with the catchment communities [30, 31] and drainage engineers at Yarra Ranges or Knox City councils [27]. For the two urban, control catchments (Br and Fe, Fig 1), we assumed that 90% of impervious surfaces in areas serviced by stormwater drainage pipes were connected (the approximate rate of connection in comparably serviced areas of L1). We outline the method of determining connection in detail in S2 Text.

SCM specification and modelling. As part of ongoing engagement with the catchment communities, councils and Melbourne Water, we recorded specifications of all SCMs installed in the Little Stringybark and Dobsons Creek catchments, together with dates of construction,

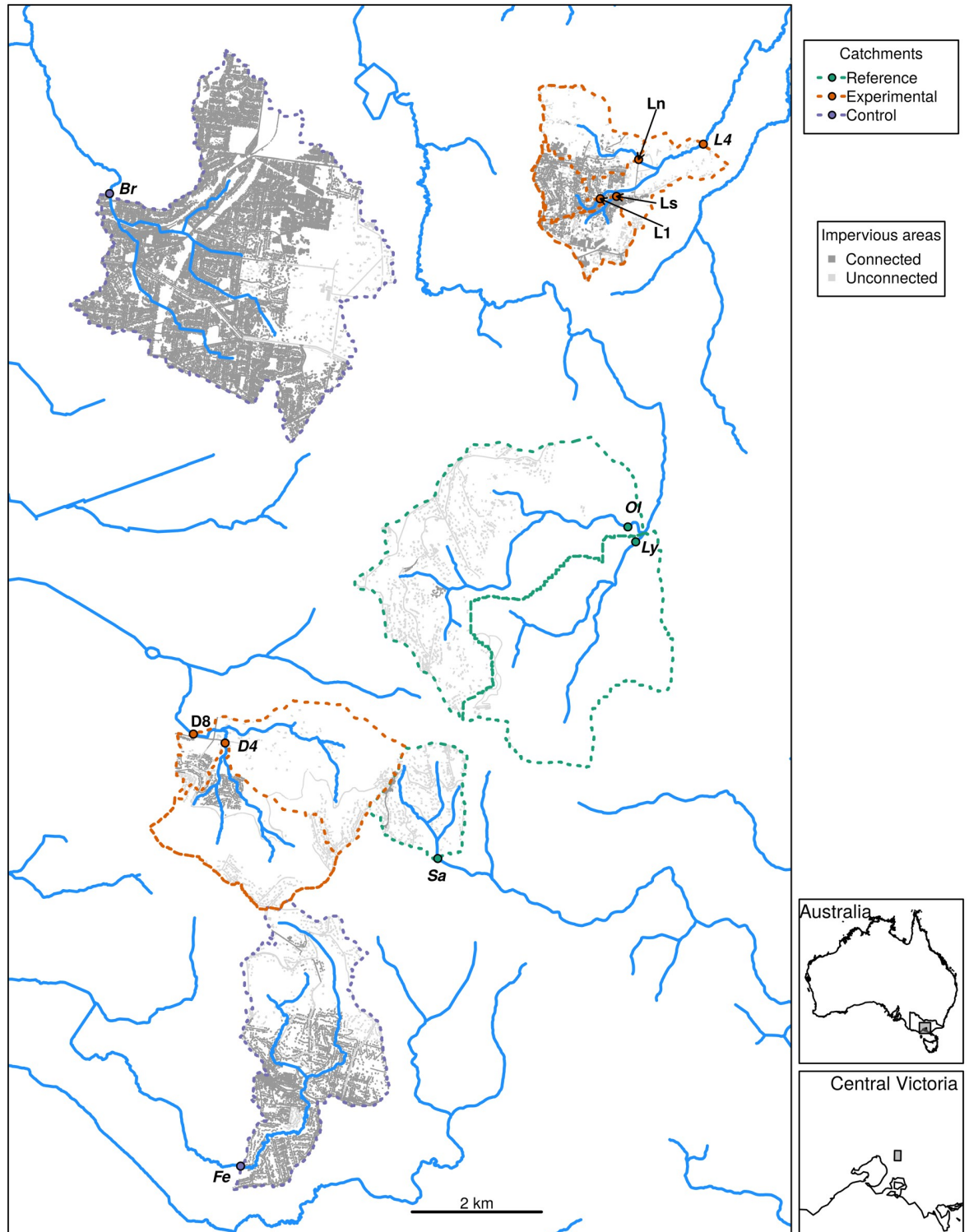


Fig 2. The 11 study sites considered in this paper and their catchments. Sites and catchment boundaries are coloured by their status as control, experimental or reference. The 7 primary catchments of the original experiment are italicized. Impervious areas in each catchment are shaded by their connection to the stormwater drainage system in 2019. (Impervious areas outside catchments are not shown.) Inset maps show the location of the main map (grey rectangle) in Victoria, Australia. (Source: Map created through data and code available at OSF: <https://osf.io/57azq/>).

<https://doi.org/10.1371/journal.pwat.0000004.g002>

and where necessary, dates of decommissioning or change in specifications. These SCM data, together with the impervious coverage data, were structured in a spatial database (see [S3 Text](#) for details).

We adapted the R [32] functions of the ‘environmental benefit’ calculator, which had been used to assess and fund SCMs in the Little Stringybark catchment (<https://tools.thewerg.unimelb.edu.au/EBcalc/>, [23]), to model water fluxes through each SCM. We wrote a series of functions to iteratively calculate a time series of EI_S for each study site over the study period (see [S4 Text](#) for details). All dates on which changes (construction or demolition of impervious surfaces, commission or decommission of SCMs) occurred in the catchment were identified. For each such date, catchment IA was partitioned into those surfaces draining directly via stormwater drainage to the stream (and S therefore = 1), and the IA draining to each terminal SCM (i.e. an SCM draining to the stormwater drainage network rather than to another SCM). S (and its sub-indices) was calculated for each SCM, accounting for retention and treatment by any upstream SCMs. From this inventory of IA and S values EI_S was calculated for the catchment by [Eq 1](#). This process was repeated for each date of change to derive a time series of EI_S .

The target volumes used to calculate S_F and S_V were described by Walsh, Fletcher [33] as ranges of volume to filter and volumes to harvest, which vary as a function of mean annual rainfall. We calculated the target ranges for the experimental catchments, with a mean annual rainfall of 956 mm, and compared the respective volumes harvested/lost and filtered for each catchment. We modelled these volumes given the arrangement of SCMs in each catchment at the end of the study period, assuming rainfall pattern of the average year that was used to calculate S (1965–1966).

Statistical model

The effect of the SCMs implemented during the project on reaches of receiving streams can be modelled by separating the putative degrading effect of urban stormwater runoff—which we term *degrd*, as measured by $\log_{10}(EI_{S1} + 0.1)$ —, from the putative restorative effect of SCMs, *restr*, as measured by ΔEI_S ; the difference between $\log_{10}(EI_{S1} + 0.1)$ and $\log_{10}(EI_S + 0.1)$. Given:

1. a set of sites that span the range of *degrd* from reference to control condition,
2. experimental sites with beginning condition similar to the control sites, with *restr* = zero before, and increasingly negative *restr* after the commencement of SCM implementation, and
3. samples of a response variable collected repeatedly at each site for a period before and after the commencement of SCM implementation,

a model that includes *degrd*, *restr*, a time effect, and a random site effect should suffice to detect any response to reduced *restr*. To illustrate such a model, we simulated two response variables, one (y) with no response to SCM implementation, and a second ($y1$) for which SCM implementation removed the degrading effect of stormwater runoff (such that the effect of *restr* equalled that of *degrd*). Variation in y and $y1$ were based on median $\log_{10}(\text{FRP})$ ([Fig 1A](#)), each with the same positive linear relationship with EI_{S1} (slope coefficient = 0.284). $y1$ was similarly positively related *restr* (same slope as EI_{S1}), while y was uncorrelated with *restr*. To illustrate the robustness of the model to external influences over the study period, we also added a linear trend over time to each of y and $y1$, that was steeper in sites with greater EI_{S1} (increasing the mean *degrd* effect over the study from 0.284 to 0.351). This effect was based on our observation of increasing wetness during our study period and an increased contribution of

pervious flows that was more pronounced in control streams than reference streams. Such effects over time unrelated to the experimental manipulation were observed in a similar experiment [34]. To simplify the example, we excluded L4 and D8 from the model so that all sites were independent: a full analysis including all sites will require dealing with spatial autocorrelation. In summary, if the model accurately assesses the experimental effects from other variation in the data, it should estimate a *degrd* effect of 0.351 and a *restr* effect of 0 for *y* and of 0.284 for *yI*. See S5 Text for further details and the code used to simulate the variables and compile the dataset for analysis.

The following description of the model for *y* applies equally to *yI*. We modelled each value of *y* in each sample as being drawn from a normal distribution thus:

$$y_i \sim \text{Normal}(\mu_i, \sigma) \quad (3)$$

where y_i is the response variable in the i th of 76 equally spaced samples, μ_i is the mean estimate for that sample, and σ is the standard deviation of the estimate. We modeled μ_i in response to predictors as a hierarchical linear model:

$$\mu_i = \alpha + \alpha[\text{site}_j] + \beta_D \text{degrd}_i + \beta_R \text{restr}_i + \beta_T[\text{site}_j]t_i + \beta_A \text{autoT}_i \quad (4)$$

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 (to account for the possibility that sites differed in their trends over time, separate from the *degrd* and *restr* effects); β_A represents the effect of temporal autocorrelation, *autoT*, among samples from each site; a . To derive the temporal autocorrelation variable, we followed an approach similar to that of Crase, Liedloff [35]: by calculating the residuals of the model without the autocorrelation term, and using the residual value for the preceding time in the same site. All α and β parameters were drawn from normal distributions with estimated means. We specified prior distributions of α and all β parameters as diffuse normal distributions (mean 0, standard deviation 5), and of $\alpha[\text{site}_j]$ as diffuse normal distributions (mean zero and estimated standard deviation). Prior distributions for standard deviation of $\alpha[\text{site}_j]$ and σ were half-Cauchy distributions (mean 0, scale 2.5). We drew inference from 10,800–12,800 posterior samples taken from 4 unthinned chains (2,700 per chain without *autoT* and 3,200 with *autoT*, having discarded 1,500 warm-up values of each chain) using the Markov Chain Monte Carlo sampler implemented in Stan [36].

We ensured the models met the standard diagnostic tests [36], and quantified accuracy by calculating the correlation coefficient between predicted and observed (simulated) values of *y* and *yI* for each of the models with an *autoT* term. Most importantly, we assessed if the models accurately detected the simulated *degrd* and *restr* effects for the variables *y* and *yI*: (*degrd* coefficient 0.35 for both variables, and *restr* coefficient 0.28 for *yI* and 0 for *y*). See S5 Text for further details, including all model code.

Finally, to illustrate the application of the model and BACRI experimental design to predict and portray responses to the experimental SCM implementation, we used the two models to predict *y* and *yI* at each of the four experimental sites under four scenarios:

- a). with EI_S as it would have been at the end of the study, had no stormwater control measures been installed (= EI_{SI}).
- b). with EI_S as it was at the end of the study, after stormwater control measures were installed. By contrasting this prediction with a), the effect of the experimental manipulation can be assessed.

- c). with *degrd* reduced by -1 (i.e. EI_S reduced by a factor a 10), to assess potential change from b) with additional SCM implementation.
- d). With zero EI_S (*degrd* set to its minimum value, *restr* set to 0), in the absence of urban stormwater impacts. Contrasting this prediction to the a), b) and c), permits an assessment of the extent to which the experimental SCMs restored γ and γI to reference levels.

Results

Spatial and temporal patterns of EI variants

1,029 SCMs in 638 projects were installed in the experimental catchments over the study period (Table 1). SCM density was similar across the sub-catchments of L4, averaging 7.8 SCM projects per ha of effective impervious area (EIA, Table 1, Fig 3A). In D8 and its sub-catchment D4, where greater focus was applied to household stormwater control, there were 18.4 projects per ha of EIA (Table 1, Fig 3A). Larger-scale SCMs implemented in Ln and Ls resulted in runoff from 96% and 92% of EIA, respectively, draining to at least one SCM, while in L1 and D4, with less space available for large-scale SCMs, less than half of EIA drained to an SCM (Fig 3B).

Volumes of impervious runoff filtered or exfiltrated by SCMs were less than the target range in all sub-catchments except Ln and D8 (Fig 3C, with pink polygon representing the target range). This was primarily because, except for these two catchments, insufficient EIA drained to SCMs (Fig 3B), even though filtered volumes tended to be within or close to the target range for individual SCM projects. In all sub-catchments, the volume of stormwater harvested or lost to evapo-transpiration was substantially lower than the target range (Fig 3D), primarily because we were unable to find sufficient harvesting demand for the water.

Despite similar SCM density (Fig 3A) and effectiveness (Fig 3C and 3D) in each of the study catchments, the reduction in EI (on a log-scale, which is most appropriate for predicting stream responses, Fig 1) varied strongly between catchments. Fig 4 shows the trend in 7 variants of EI over the study period in each of the 6 experimental catchments. EI_{SI} , which assumes no effect of SCMs, increased slightly in all catchments (with urban growth). All other variants, which account for SCM effects, reduced over time compared to EI_{SI} .

Likely responses to SCM implementation among the experimental catchments are more dependent on initial EI and catchment size than on the extent of SCM implementation in the catchment. For instance, the treatment of 4.4 ha of the 17 ha of EIA in L1 (Table 1) resulted in very little change in EI, even by the broadest definition of disconnection (EI_{S0} , Fig 4D), while treatment of 8.8 ha of 8.9 ha in the similarly sized Ln catchment resulted in the largest decline

Table 1. Summary statistics for stormwater control in the 6 experimental sub-catchments.

Stormwater control	D4	D8	Ln	Ls	L1	L4
Number of SCMs	263	451	144	122	252	578
Number of SCM projects	161	266	80	87	162	372
EIA (ha)	9.0	16.3	9.2	12.9	17.0	45.6
Treated EIA (ha)	3.6	9.9	8.9	11.9	4.4	27.6
Runoff volume harvested/lost (ML/y)	9.0	17.0	15.0	20.0	9.0	49.0
Runoff volume filtered (ML/y)	4.0	33.0	21.0	13.0	7.0	45.0

Stormwater control measure (SCM) projects were commissioned as single projects and consisted of one or more SCMs. Note that the counts for D8 include those listed for its sub-catchment D4, and those for L4 include those for its sub-catchments L1, Ls and Ln. Treated EIA is the effective impervious area draining to one or more SCMs. The estimates of runoff volume harvested or lost, and volume filtered, are modelled volumes given the arrangement of SCMs in each catchment at the end of the study period assuming rainfall pattern of the average year that was used to calculate S (1965–1966).

<https://doi.org/10.1371/journal.pwat.0000004.t001>

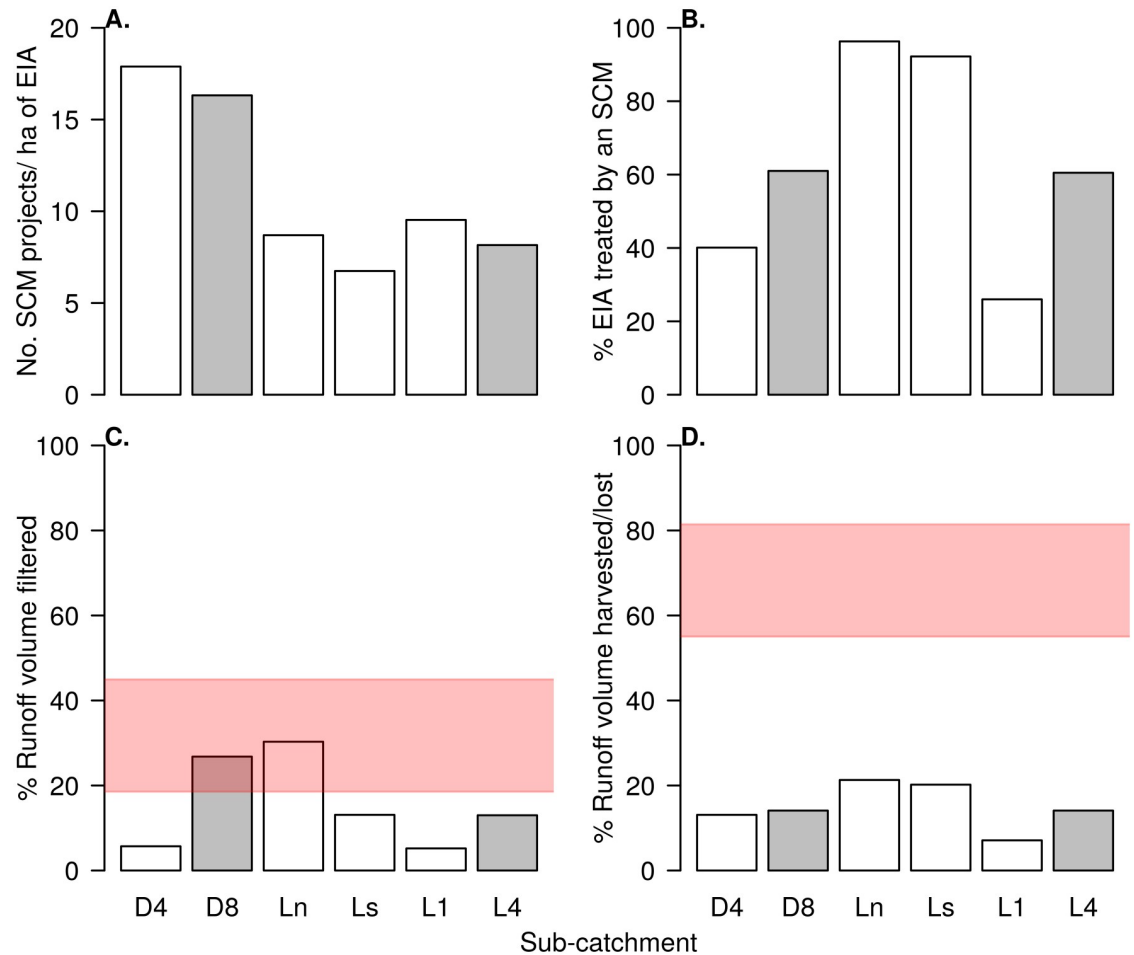


Fig 3. Stormwater control spread and effectiveness in each of the 6 experimental sub-catchments. (A) Number of SCM projects per ha of effective impervious area (EIA). (B) Percentage of EIA that drains to at least one SCM. (C) Percentage of impervious runoff harvested or lost to evapo-transpiration in SCMs. (D) Percentage of impervious runoff filtered through SCMs or to soil. In (C) and (D), the pink polygons indicate the target percentages proposed by Walsh, Fletcher [33]. D8 and L4 bars are shaded to indicate that these sub-catchments contain other sub-catchments. D4 is part of D8, and L1, Ln and Ls are part of L4; see Fig 2.

<https://doi.org/10.1371/journal.pwat.0000004.g003>

in all variants of EI (Fig 4C), while treatment of 3.7 of the 8.9 ha of EIA in D4 (catchment area ~4 times larger than L1 and Ln) resulted in only minor declines in EI (Fig 4A and 4B).

EI_{S0} indicates the extent to which impervious areas drain to any sort of SCM, while EI_{S1} indicates growth in EI in the absence of SCMs. In all cases, EI variants accounting for the performance of SCMs falls between these two extremes. In most catchments EI_{SW} was the closest to EI_{S0} , suggesting good performance in water quality treatment. [The model assumes optimal water quality for exfiltrated flows, and high quality of flows through filtration systems, based on reported performance: 37, 38]. EI_{SV} declined the least in all sub-catchments, indicative of our inability to reduce flow volumes adequately. EI_{SF} and EI_{SR} fell between the two extremes, suggesting moderate success of installed SCMs in restoring filtered flows and in reducing the frequency of uncontrolled runoff. EIs, as the average of the four sub-metrics, was usually of a similar value to EI_{SF} and EI_{SR} .

The extent to which these observed declines in EI match differences between control and reference sites can be best illustrated and interpreted by separating the degrading (*degrad*) effects of urban growth and urban stormwater runoff in each site, as indicated by EI_{S1} (Fig 5A), from

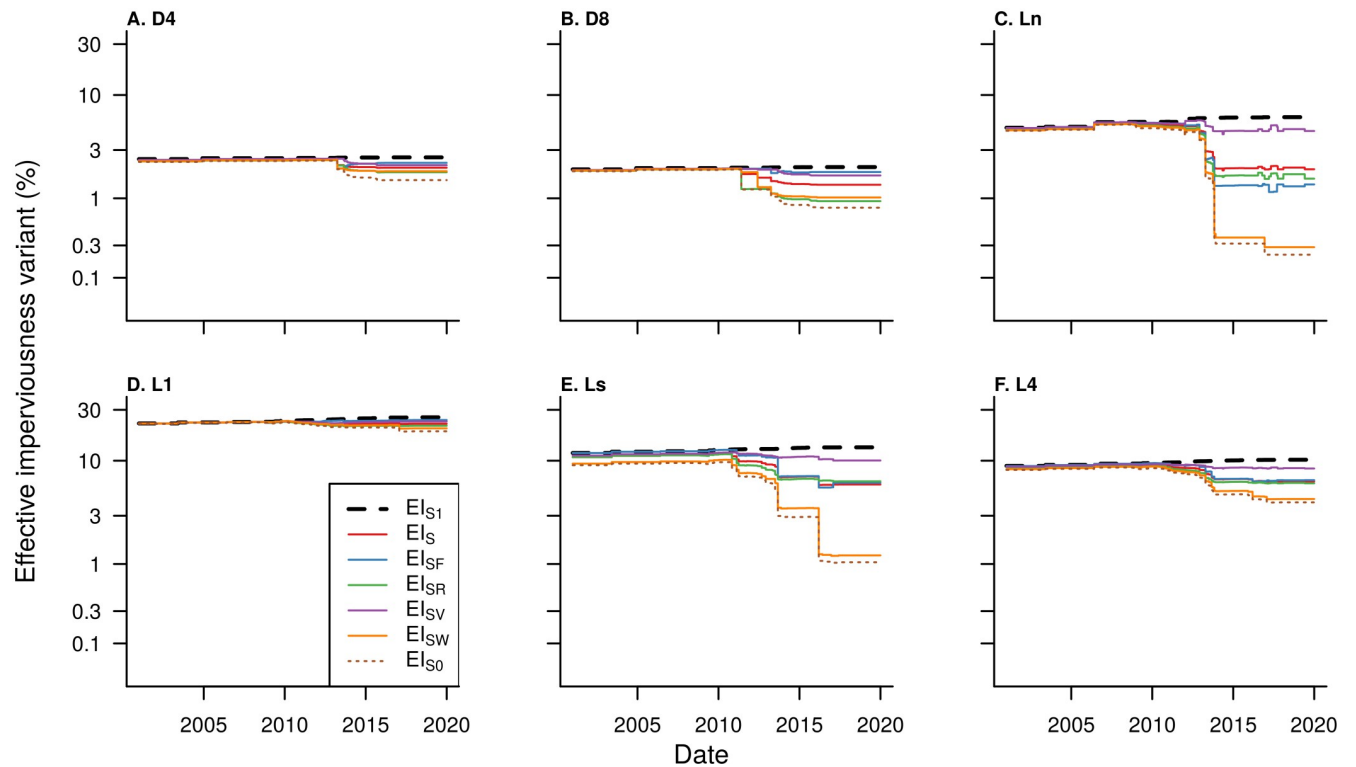


Fig 4. Time series of effective imperviousness (EI) variants achieved in each of the experimental catchments. Subscripts denote variants. S1: all effective impervious areas upstream of stormwater control measures (SCMs) unweighted (assuming SCMs have no effect, $S = 1$). S0: all effective impervious areas upstream of SCMs set to 0 (assuming all SCMs retain stormwater perfectly). For other subscripts, all effective impervious surfaces upstream of SCMs are weighted by the stormwater stream impact metric (S), or one of its sub-indices, of the most downstream SCM in any train of SCMs. SF: the filtered-flow sub-metric. SR: the runoff frequency sub-metric. SV: the volume reduction sub-metric. SW: the water quality sub-metric.

<https://doi.org/10.1371/journal.pwat.0000004.g004>

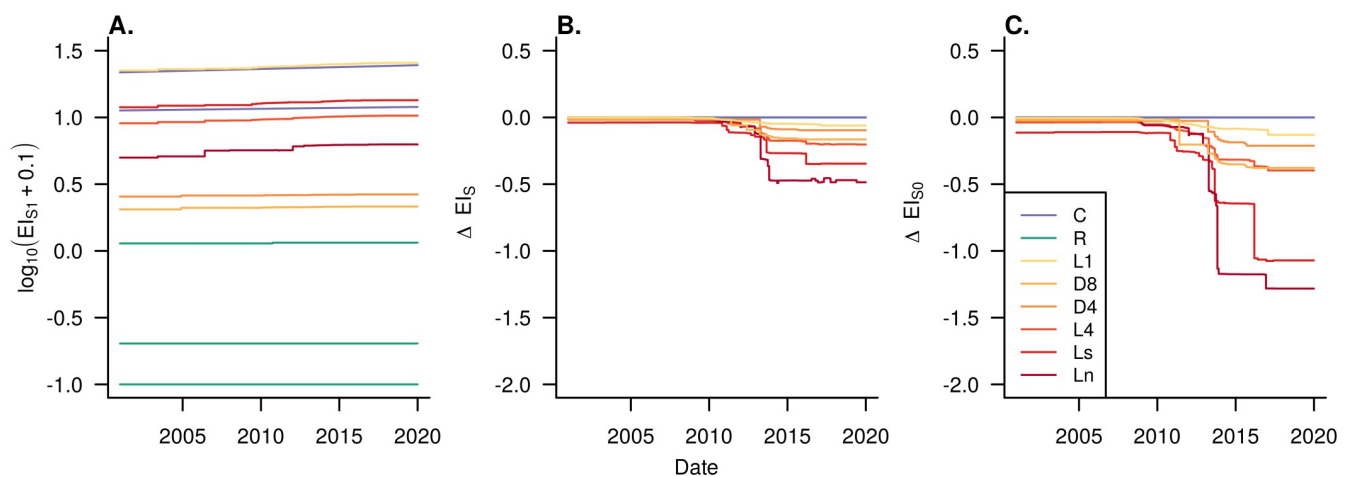


Fig 5. Time series of predictors of effects of urban stormwater runoff and control in the 11 study catchments. (A) *degrd*: $\log_{10}(x + 0.1)$ -transformed effective imperviousness assuming no stormwater control (EI_{S1}). (B) *restr*, or ΔEI_S , the change in $\log_{10}(x + 0.1)$ -transformed EI weighted by the stormwater stream impact metric, S . (C) ΔEI_{S0} , the change in $\log_{10}(x + 0.1)$ -transformed EI with all impervious surfaces draining to an SCM weighted zero for the two control (C) sub-catchments, three reference (R) sub-catchments and the six experimental sub-catchments, indicated by code (see Fig 2). Note all control and reference values in panels B and C equal 0.

<https://doi.org/10.1371/journal.pwat.0000004.g005>

the restorative (*restr*) effect of installed SCMs, which can be indicated by the (log-) differences between EI_S and EI_{S1} , ΔEI_S (Fig 5B). Similarly, ΔEI_{S0} can be calculated by the (log-) differences between EI_{S1} and EI_{S0} (Fig 5C).

EI_{S1} in the 6 experimental catchments ranged from the high of L1, with similar EI_{S1} to the most urban control (Br) to the low of D8, with EI_{S1} just higher than the most urban reference site, Sa (Fig 5A). Although there was a small amount of growth in EI_{S1} over the study period, it was small compared to the intersite variation in EI_{S1} . The effect of SCMs in the experimental catchments as measured by reduction in EI_S varied from near zero to half an order of magnitude in the case of Ln (Fig 5B). Reductions in EI_{S0} were larger again (Fig 5C).

Statistical model for assessing stream response

The models of y and yI predicted the observed (simulated) data well ($R = 0.97$ for both models, see S5 Text), affording confidence in the models' estimates of treatment effects. The estimated *degrd* effect was 0.35 (95% credible intervals 0.34, 0.36) in both models, equivalent to the known mean effect in the simulated data. The estimates for the *restr* effects matched the known effects for the simulated data: for y , the mean estimated *restr* coefficient was 0.01 (95% credible intervals -0.05, 0.07), which did not differ from the simulated *restr* effect of 0, and for yI , the mean estimated *restr* coefficient was 0.29 (95% credible intervals 0.24, 0.36), which did not differ from the simulated *restr* effect of 0.28 (Fig 6A and 6C). The experimental design and model structure thus permit accurate detection of changes in response to *restr* in experimental sites, even in the context of varying temporal trends among the study sites.

Fig 6B and 6D show model predictions for y and yI , respectively, under the four scenarios that aim to illustrate responses to the SCM implementation in experimental catchments, relative to predicted levels in control and reference sites.

The degree of stormwater control achieved in the experiment, as estimated by *restr*, was sufficient to detect a decrease in yI in Ln, Ls, and perhaps D4, but not in L1 (red vs orange points in Fig 6D). The model also predicted that, even if EI_S could be reduced ten-fold (i.e. *restr* (ΔEI_S) = -1), yI would remain somewhat higher than the reference state in all cases (light green vs dark green points in Fig 6D).

Discussion

The stream stormwater impact metric, S , advances on the dominant approach to estimating SCM performance by measures of pollutant load reduction compared to untreated stormwater runoff [e.g. 11–13]. Its use in EI_S advances the assessment of catchment-scale effects of SCMs, permitting a more informative assessment than previous studies, which have assumed a uniform effect of dispersed SCMs across catchments [24, 39]. The use of EI variants as continuous predictor variables to assess the effects of our experiment advances categorical approaches to before-after-control-impact assessment of dispersed SCMs [e.g. 24].

The formulation of S is consistent with the principles of stormwater management for stream protection proposed by Walsh, Booth [14]. By scaling each of its sub-metrics between untreated stormwater runoff and a reference or target condition, it accounts for natural variation in stream function across regions. This attribute of S meets the first principle: the setting of a target state for the stream to be protected. The filtered flow and water quality sub-metrics address the third principle: to deliver flows in a quality and flow regime that mimic predevelopment hydrologic processes. The runoff frequency sub-metric addresses the fourth principle: SCMs should prevent untreated flows to streams in all but rare, large storms. The volume reduction sub-metric (in concert with the other sub-metrics) addresses the second principle: to mimic the predevelopment water balance.

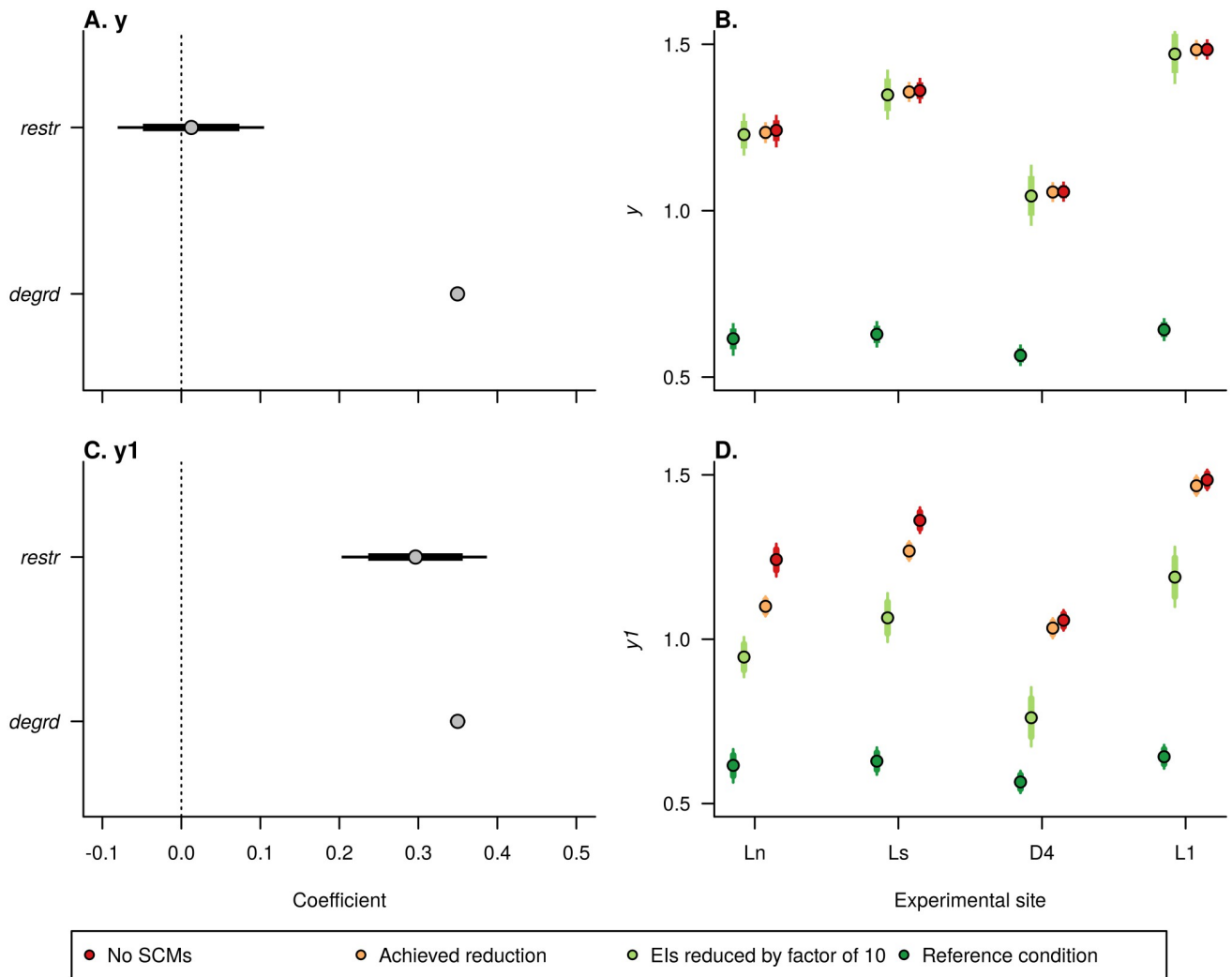


Fig 6. Results from the hierarchical linear model of two simulated response variables to urban stormwater runoff and control. (A, B) show results for y , simulated to be unaffected by stormwater control (as indicated by *restr*, while (C, D) show results for $y1$, simulated to be equally affected by stormwater runoff (as indicated by *degrd*, as it was at the start of the study) and stormwater control. (A, C) show coefficient estimates for each effect, showing that the model correctly estimated these effects. (The *degrd* effect on both variables was simulated to increase over time, resulting in the mean *degrd* effect for $y1$ being greater than the *restr* effect). (B, D) show model predictions under four treatment scenarios at the end of the study: as they would have been without any stormwater control (red), the *restr* effect (ΔEI_S reduction) actually achieved (orange), a change in ΔEI_S of -1 (i.e. a 10-fold reduction in EI_S , light green), and reducing EI_S to zero (reference condition, dark green). Each estimate is illustrated by the median (point), 80th (thick error bars) and 95th (thin error bars) percentile credible intervals.

<https://doi.org/10.1371/journal.pwat.0000004.g006>

The fifth principle of Walsh, Booth [14]—that SCMs should be applied to all impervious surfaces in the catchment of the target stream—extends beyond the performance of individual SCMs. It was founded on the observation that small areas of connected impervious surfaces in a catchment potentially have a large influence on stream degradation, as evidenced by the non-linear response of stream ecosystems to EI (Fig 1). EI_S provides a link between SCM performance and responses of stream ecosystems at the catchment scale. The small reductions in EI_S in catchments where we were unable to intercept runoff from most impervious surfaces demonstrate the importance and challenge of this principle. It was only in Ln, the catchment in which we were able to intercept runoff from almost 100% of impervious surfaces in the catchment that substantial reduction in EI_S was achieved.

Our aim in designing and implementing SCMs in the experimental catchments was to maximise environmental benefit [sensu 15], and thus minimise S , given a limited budget, and other limitations arising from working in an established suburban catchment. The resulting reductions in EI_S were small in most sub-catchments, indicating that the ultimate goal of the experiment—to reduce stormwater runoff impacts sufficiently to restore the receiving creeks—has not yet been achieved. In the following section we discuss reasons for target reductions in S not being fully realised. We then consider the form of statistical model required to assess stream responses to the protracted and variable changes in EI_S that were achieved in this and similar experiments, which has potentially broader applications to studies of similar design.

Assessing the success of the experimental intervention using patterns in the EI variants

The area of impervious surfaces draining to SCMs was substantially less than our objective [principle 1 of 14] in all experimental catchments except Ln and Ls (as indicated by EI_{So} , Fig 3). An optimal outcome would have been to fund as many dispersed treatments as possible in the catchment, with a final SCM near the outlet of pipes opening to the stream, to permit treatment of runoff from missed impervious surfaces, and to improve the cumulative performance of upstream SCMs.

Despite several campaigns of community engagement, which overcame many barriers to community participation, SCMs were installed on only 28% of the private properties with drainage connection in the L4 catchment [26; but this number grew to 30% in the following 4 years, because of a planning provision placed on the catchment, 40]. Two engagement rounds in the D8 catchment achieved SCMs on 46% of the connected properties. Treatment of every property in catchments such as ours is unlikely to be achievable because there will always be a limit to community engagement and acceptance. Furthermore, appropriate space for SCMs was not available to treat all impervious surfaces on many of the properties that registered interest in the project. Similarly, although 100s of dispersed raingardens were installed to capture runoff from roads of the experimental catchments, space was limited in many cases, leaving large areas of road untreated. In most cases, there will be limits to the ability of dispersed SCMs to meet the objective of intercepting runoff from all impervious surfaces. We thus aimed to also install, wherever possible, SCMs at the end of pipes, near where they enter the stream. However, we were hampered by a lack of space near pipe outlets. Even in Ln and Ls, space at the end of several large pipes was limited and we resorted to SCMs that captured and filtered little more than the first 1–2 mm of runoff.

Ultimately the lack of space was a function of past planning decisions that did not reserve public space along drainage lines. Indeed, in much of the catchment, many large council stormwater drainage pipes, and the streams themselves, flow through private land. As a result, many of the large end-of-pipe SCMs that we installed were on private land, with the cooperation of owners [26, 27]. We posit that the limiting space we encountered in our experimental catchments, both in upland areas and at the end of pipes, is likely to be typical of many suburban landscapes, posing a significant challenge for stormwater retrofits aiming to restore degraded streams. Even with the long-term growth in SCM numbers that is likely as a result of the planning provision for the Little Stringybark Creek catchment [40], the lack of space at the end of pipes for final treatments will remain a barrier to restoration without the strategic purchase of small areas of low-lying land for SCMs.

The SCMs that were installed as part of the experiment varied widely in performance (range of S 0.01–0.89 in L4, and 0.30–0.96 in D4), underlining the utility of EI_S for ensuring a

comparable measure of experimental effect within our study, and with other studies. The temporal and inter-site patterns of ΔEI_S (Fig 5B), weighted by SCM performance, differed substantially from ΔEI_{S0} (Fig 5C), which weights all SCMs uniformly, as did Roy, Rhea [24] and Loperfido, Noe [39]. Optimal performance of individual SCMs, as measured by S , required a combination of harvesting and loss of at least 55% of impervious runoff (Fig 3D), and retaining and filtering of 19–45% (Fig 3C) to achieve a low frequency of uncontrolled flow to the stream. Some infiltration systems performed poorly for EI_{SF} because they filtered too much water (61 systems in L4, 7 in D8). However, at the catchment scale, infiltration volumes were lower than optimal because of the lack of SCM coverage, and the local excess infiltration from these SCMs improved performance at the catchment-scale. However, to achieve near-full SCM coverage without exceeding infiltration targets, substantially greater harvesting and loss of impervious runoff than we achieved will be required to avoid more widespread infiltration excess.

Water demand was enough to meet the volume reduction target of 55% in 157 SCM projects in L4 and 34 in D8. All of these projects harvested runoff for domestic uses from relatively small areas: roofs of $< 625 \text{ m}^2$ (80% $< 270 \text{ m}^2$). Two projects harvested water for large demands from larger areas, but were of limited influence at the catchment scale. Tanks and treatment systems receiving runoff from the entire 0.33 ha of impervious area of the catchment's petrol station provide 96% of its car wash's 1.8 ML/y demand. Despite this large demand, the harvested water only accounted for 50% of the runoff generated by the station's impervious area, and S for this project was 0.15. The influence of this high-performing project was ultimately small, as the petrol station's impervious area constituted only 2.5% of the effective impervious surfaces in the Ls catchment. The largest-scale harvesting scheme collected runoff from a large educational precinct in the Ln catchment with 2.9 ha of impervious surfaces providing 58% of the toilet-flushing demands of the precinct's schools, and the dry-weather irrigation demand of its playing fields, which totalled $\sim 6 \text{ ML/y}$. However, this large harvesting volume, constituted only 21% of the runoff generated by its impervious catchment. Despite some SCM projects achieving the volume reduction target, generally roof runoff on private properties produced more water than the property's demand. Furthermore, the lack of SCM coverage meant no runoff was harvested from 65% of the L4 catchment's roofs, and very little from paved surfaces and roads, which make up 41% of the catchments' impervious area.

The opportunity for augmentation of water supply to protect streams such as our experimental streams is large [33], but without an urban water management policy that prioritises harvesting of stormwater, stream protection remains challenging. Despite these shortcomings of the experimental implementation of SCMs, the range of reductions of EI_S among the experimental catchments provides a framework with potential for assessing the efficacy of stormwater control for stream protection.

Statistical model for BACRI experiments

The trial of simulated variables with known contrasting responses to reduction in EI_S confirmed the effectiveness of the proposed model to assess stream response to SCM implementation. It also confirmed that the level of control achieved in the experiment should be sufficient to detect responses, at least if the magnitude of responses are close to the magnitude of degradation from stormwater runoff, and sample replication is similar to our simulation. It is unlikely that any variation over time unrelated to changes in EI_S in our study sites would be as pronounced as the linear increases we simulated in γ and γ_1 . If change over time unrelated to change in EI_S was small, then sensitivity of analyses for real response data is likely to be greater. The temporal autocorrelation term ($autoT$, Eq 3) was not different from zero for our simulated data, but it is feasible that real response data may have less of a time effect and more of a

temporal autocorrelation effect. Some response variables are likely to warrant further predictor variables in the model. For instance, contaminant concentrations and the ability of SCMs to reduce stormwater concentrations are both likely to be a function of antecedent rainfall, suggesting the addition of a rainfall variable is warranted, together with interactions between *degrd* and rainfall, and *restr* and rainfall.

Restoration trajectories of degraded ecosystem are highly uncertain [41], but the reverse of the degradation trajectory provides the most optimistic expectation of any restoration effort. If we reduce EI_S through implementation of SCMs, we would, at best, expect to see a 1:1 reversal of the degradation trend, as modelled by the simulated variable $y1$ (Fig 6C). If the restoration response (estimated by the *restr* coefficient in the model) is smaller than the degradation response (the *degrd* coefficient), the model should still detect a response, but with a smaller coefficient for *restr* than for *degrd* (this was the case in our simulated data, but it was because we increased the *degrd* effect over time). It is also feasible that a discrepancy between coefficients for *restr* and *degrd* could mean that the averaging of the four sub-metrics comprising S does not adequately represent the impact of urban stormwater runoff. Such a possibility could be explored by investigating responses to the variants of EI_S (Fig 3). Lags in response to restoration are also likely [41]. Such lags could be explored through lagging the *restr* predictor.

The model structure we have used provides an analysis analogous to the various conventional approaches to Before-After-Control-Impact analyses [see review by 25]. The categorical 'Control-Impact(-Reference)' effect is equivalent to the continuous *degrd* effect, as a measure of the putative impact driving differences between catchments. The interaction between the 'Before-After' effect and the 'Control-Impact(-Reference)' effect, which provides the inference of an 'impact' effect is equivalent to the *restr* effect, because the only cases when it is non-zero are experimental ('Impact') catchments at times after experimental SCM installation has begun. An 'Impact' effect is thus inferred by a *restr* effect in the same direction as the *degrd* effect. This statistical approach using continuous measures of impact is likely to have broader application in assessing other protracted impacts or management interventions such as the effects of re-vegetation.

Conclusion

In this paper, we described the design and implementation of a catchment-scale BACRI experiment testing the effects on stream ecosystems of dispersed urban stormwater control in 6 experimental catchments. To assess the expected effect of SCMs installed in the experimental catchments, we developed a new variant of EI, weighted by S , a performance metric for SCMs. By integrating a catchment scale predictor of stream response with a measure of SCM performance, EI_S has potential broader application for assessing and predicting the cumulative effects of SCMs at catchment-scale across regions.

We also proposed, tested and validated a modelling approach for assessing BACRI experiments such as ours, which have necessarily protracted and varied degrees of intervention in the experimental sites. Validating the model demonstrated that our experimental design and manipulation were likely to be sufficient to detect stream responses to the SCMs implemented in the project. However, the performance and spatial extent of SCMs installed were insufficient to meet the ultimate goal of the experiment of restoring the streams of the experimental catchments.

While the implementation of SCMs across our experimental catchments did not achieve the reduction in EI_S that was originally aimed, the barriers we encountered provide important lessons for how urban planning and urban water management need to change for the protection and restoration of stream ecosystems. Stormwater control is likely to require control

measures at multiple scales, including at the scale of individual properties (e.g. rainwater tanks, raingardens), on streetscapes (e.g. street-tree pits, raingardens), and at the end of pipes (e.g. large-scale harvesting systems, biofiltration systems, wetlands) as a final treatment before receiving streams. As a result, the reservation of (at least) small areas along flow lines as public open space for stormwater interception at the planning stage is required. Stormwater control measures will provide greater protection to stream ecosystems if they mimic the pre-development water balance [14]. To achieve such an aim, substantial demand for stormwater is required, and such demands are only likely to be met by urban water governance that places a priority on stormwater harvesting and use.

Despite the project's SCM implementation being insufficient to elicit complete restoration of the receiving streams, our statistical modelling of simulated data suggests that the experimental intervention was likely to be large enough to detect in-stream responses, and therefore enough to test the broader hypothesis that impervious runoff delivered through conventional drainage systems was the primary degrading process requiring remediation to permit restoration of the receiving stream ecosystem. We have monitored hydrologic, water quality and ecological variables in the receiving streams over the period of study, and future studies will report on those in-stream responses.

Supporting information

S1 Text. Derivation and formulation of the stream stormwater impact metric, S.
(PDF)

S2 Text. Impervious surface mapping methods.
(PDF)

S3 Text. Stormwater control measure specification and location data.
(PDF)

S4 Text. Calculation of effective imperviousness variants and water fluxes in the study sub-catchments.
(PDF)

S5 Text. Statistical model for assessing stream responses to stormwater control measures.
(PDF)

Acknowledgments

We sincerely thank the following who have provided support and encouragement in undertaking the project: from Melbourne Water, Graham Rooney, Rhys Coleman, Toby Prosser, Sharyn Rossrakesh, Marion Urrutiaguer, Chris Chesterfield, Rachele Adamowicz, Michael Godfrey, Leon Harvey, Justin Lewis, Peter Morison, Kelly Bayton, Andrew Camenzuli, Bruce Rhodes; from Yarra Ranges Council, Beth Wallis, Caroline Carvalho, Vjekoslav Matic, Stephanie Hamel, Patrick Jeschke, Claudette Fahy, Mark Varmalis, Ashley Tenson, Darryl Cockett, Gary Whitehead; from Knox City Council, Daniella Gerente, Caroline Carvalho, Matt Potter; Ann Allworth (Victorian Department of Environment, Land, Water and Planning), Francis Pamminger (Yarra Valley Water), Lindsey Brown, Tommy Plahcinski, Jamie Taunton, Geoff Vietz, Kathy Russell, Tony Ladson, Barry Hart, Helen Brown, Marit Larson, Perrine Hamel, Julia White, Rhiannon Birch, Jemima Milkins, Belinda Hatt, Slobodanka Stojkovic, Simon Roberts, John Beardall, Sally Boer, Monica Tewman, Ashley Macqueen, Tony Lovell, Mike Sammonds, Sara Lloyd, Pascal Horton, Carleen Mitchell, Rob James, Genevieve Hehir, Lee Hazel, the communities of the Little Stringybark Creek and Dobsons Creek catchments, and

Raj Rosrakesh. We thank Mike Sammonds, Rob James, Congying Li, Jeremie Bonneau, Tim Craig, and Hugh Duncan for contributions to an early draft of the manuscript.

Author Contributions

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

Data curation: Christopher J. Walsh, Darren G. Bos, Peter Poelsma, Joshphar Kunapo, Moss Imberger.

Formal analysis: Christopher J. Walsh, Matthew J. Burns, Peter Poelsma, Joshphar Kunapo, Moss Imberger.

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

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

Methodology: Christopher J. Walsh, Tim D. Fletcher, Peter Poelsma, Joshphar Kunapo.

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

Software: Christopher J. Walsh, Matthew J. Burns.

Validation: Christopher J. Walsh, Matthew J. Burns, Moss Imberger.

Visualization: Christopher J. Walsh.

Writing – original draft: Christopher J. Walsh, Matthew J. Burns, Peter Poelsma, Joshphar Kunapo.

Writing – review & editing: Christopher J. Walsh, Matthew J. Burns, Tim D. Fletcher, Darren G. Bos, Peter Poelsma, Joshphar Kunapo, Moss Imberger.

References

1. Walsh CJ, Roy AH, Feminella JW, Cottingham PD, Groffman PM, Morgan RP. The urban stream syndrome: current knowledge and the search for a cure. *J N Am Benthol Soc.* 2005; 24(3):706–23. <https://doi.org/10.1899/04-028.1>
2. Vietz GJ, Sammonds MJ, Fletcher TD, Walsh CJ, Rutherford ID, Stewardson MJ. Ecologically relevant geomorphic attributes of streams are impaired by even low levels of watershed effective imperviousness. *Geomorphology.* 2014; 206(1):67–78. <https://doi.org/10.1016/j.geomorph.2013.09.019>
3. Hawley RJ, Vietz GJ. Addressing the urban stream disturbance regime. *Freshwater Sci.* 2016; 35(1):278–92. <https://doi.org/10.1086/684647>
4. Hatt BE, Fletcher TD, Walsh CJ, Taylor SL. The influence of urban density and drainage infrastructure on the concentrations and loads of pollutants in small streams. *Environ Manage.* 2004; 34(1):112–24. <https://doi.org/10.1007/s00267-004-0221-8> PMID: 15383877
5. McGrane SJ. Impacts of urbanisation on hydrological and water quality dynamics, and urban water management: a review. *Hydrol Sci J.* 2016; 61(13):2295–311. <https://doi.org/10.1080/02626667.2015.1128084>
6. Kaushal SS, Wood KL, Galella JG, Gion AM, Haq S, Goodling PJ, et al. Making 'chemical cocktails'—Evolution of urban geochemical processes across the Periodic Table of elements. *Appl Geochem.* 2020:104632. <https://doi.org/10.1016/j.apgeochem.2020.104632> PMID: 33746355
7. Russell KL, Vietz GJ, Fletcher TD. How urban stormwater regimes drive geomorphic degradation of receiving streams. *Prog Phys Geog.* 2020; 44(5):746–78. <https://doi.org/10.1177/0309133319893927>
8. Fletcher TD, Shuster W, Hunt WF, Ashley R, Butler D, Arthur S, et al. SUDS, LID, BMPs, WSUD and more—the evolution and application of terminology surrounding urban drainage. *Urban Water J.* 2015; 12(7):525–42. <https://doi.org/10.1080/1573062X.2014.916314>
9. King County. *Surface Water Design Manual.* Seattle, WA: King County Department of Public Works, Surface Water Management Division, 1990.

10. Victorian Stormwater Committee. Urban stormwater: best practice environmental management guidelines. Melbourne: CSIRO; 1999. 268 p.
11. Clary J, Urbonas B, Jones J, Strecker E, Quigley M, O'Brien J. Developing, evaluating and maintaining a standardized stormwater BMP effectiveness database. *Water Sci Technol*. 2002; 45(7):65–73. <https://doi.org/10.2166/wst.2002.0118> PMID: 11989894
12. Liu Y, Cibin R, Bralts VF, Chaubey I, Bowling LC, Engel BA. Optimal selection and placement of BMPs and LID practices with a rainfall-runoff model. *Environ Model Softw*. 2016; 80:281–96. <https://doi.org/10.1016/j.envsoft.2016.03.005>
13. Li Q, Wang F, Yu Y, Huang Z, Li M, Guan Y. Comprehensive performance evaluation of LID practices for the sponge city construction: a case study in Guangxi, China. *J Environ Manage*. 2019; 231:10–20. <https://doi.org/10.1016/j.jenvman.2018.10.024> PMID: 30326334
14. Walsh CJ, Booth DB, Burns MJ, Fletcher TD, Hale RL, Hoang LN, et al. Principles for urban stormwater management to protect stream ecosystems. *Freshwater Sci*. 2016; 35(1):398–411. <https://doi.org/10.1086/685284>
15. Walsh CJ, Fletcher TD, Bos DG, Imberger SJ. Restoring a stream through retention of urban stormwater runoff: a catchment-scale experiment in a social-ecological system. *Freshwater Sci*. 2015; 34(3):1161–8. <https://doi.org/10.1086/682422>
16. Walsh CJ, Fletcher TD, Ladson AR. Stream restoration in urban catchments through re-designing stormwater systems: looking to the catchment to save the stream. *J N Am Benthol Soc*. 2005; 24(3):690–705. <https://doi.org/10.1899/04-020.1>
17. Walsh CJ, Fletcher TD, Vietz GJ. Variability in stream ecosystem response to urbanization: Unraveling the influences of physiography and urban land and water management. *Prog Phys Geog*. 2016; 40(5):714–31. <https://doi.org/10.1177/0309133316671626>
18. Walsh CJ, Kunapo J. The importance of upland flow paths in determining urban effects on stream ecosystems *J N Am Benthol Soc*. 2009; 28(4):977–90. <https://doi.org/10.1899/08-161.1>
19. Hale RL, Turnbull L, Earl SR, Childers DL, Grimm NB. Stormwater infrastructure controls runoff and dissolved material export from arid urban watersheds. *Ecosystems*. 2015; 18(1):62–75. <https://doi.org/10.1007/s10021-014-9812-2>
20. Baruch EM, Voss KA, Blaszcak JR, Delesantro J, Urban DL, Bernhardt ES. Not all pavements lead to streams: variation in impervious surface connectivity affects urban stream ecosystems. *Freshwater Sci*. 2018; 37(3):673–84. <https://doi.org/10.1086/699014>
21. Leopold LB. Hydrology for urban land planning—a guidebook on the hydrologic effects of urban land use. Geological Survey Circular. Washington: US Geological Survey, 1968 554.
22. Walsh CJ, Fletcher TD, Ladson AR. Retention capacity: a metric to link stream ecology and stormwater management. *J Hydrol Eng*. 2009; 14(4):399–406. [https://doi.org/10.1061/\(ASCE\)1084-0699\(2009\)14:4\(399\)](https://doi.org/10.1061/(ASCE)1084-0699(2009)14:4(399))
23. Fletcher TD, Walsh CJ, Bos D, Nemes V, RossRakesh S, Prosser T, et al. Restoration of stormwater retention capacity at the allotment-scale through a novel economic instrument. *Water Sci Technol*. 2011; 64(2):494–502. <https://doi.org/10.2166/wst.2011.184> PMID: 22097025
24. Roy AH, Rhea LK, Mayer AL, Shuster WD, Beaulieu JJ, Hopton ME, et al. How much is enough? Minimal responses of water quality and stream biota to partial retrofit stormwater management in a suburban neighborhood. *PLoS One*. 2014; 9(1):e85011. <https://doi.org/10.1371/journal.pone.0085011> PMID: 24465468
25. Downes BJ, Barmuta LA, Fairweather PG, Faith DP, Keough MJ, Lake PS, et al. Monitoring ecological impacts: concepts and practice in flowing waters. Cambridge, UK: Cambridge University Press; 2002.
26. Bos DG, Brown HL. Overcoming barriers to community participation in a catchment-scale experiment: building trust and changing behavior. *Freshwater Sci*. 2015; 34(3):1169–75. <https://doi.org/10.1086/682421>
27. Burns MJ, Wallis E, Matic V. Building capacity in low-impact drainage management through research collaboration. *Freshwater Sci*. 2015; 34(3):1176–85. <https://doi.org/10.1086/682565>
28. Prosser T, Morison PJ, Coleman RA. Integrating stormwater management to restore a stream: perspectives from a waterway management authority. *Freshwater Sci*. 2015; 34(3):1186–94.
29. Government of Victoria. State Environment Protection Policy (Waters) Victorian Government Gazette, No. S 499 23 October 2018. Melbourne, Australia: Victorian Government Printer, 2018.
30. Bos JJ, Brown RR. Governance experimentation and factors of success in socio-technical transitions in the urban water sector. *Technol Forecast Soc Change*. 2012; 79(7):1340–53. <https://doi.org/10.1016/j.techfore.2012.04.006>

31. Brown HL, Bos DG, Walsh CJ, Fletcher TD, Rossrakesh S. More than money: how multiple factors influence householder participation in at-source stormwater management. *Journal of Environmental Policy and Management*. 2016; 59:79–97. <https://doi.org/10.1080/09640568.2014.984017>
32. R Core Team. R: a language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing; 2020.
33. Walsh CJ, Fletcher TD, Burns MJ. Urban stormwater runoff: a new class of environmental flow problem. *PLoS One*. 2012; 7(9):e45814. Epub 19 September 2012. <https://doi.org/10.1371/journal.pone.0045814> PMID: 23029257
34. Shuster W, Rhea L. Catchment-scale hydrologic implications of parcel-level stormwater management (Ohio USA). *J Hydrol*. 2013; 485:177–87. <https://doi.org/10.1016/j.jhydrol.2012.10.043>
35. Crase B, Liedloff AC, Wintle BA. A new method for dealing with residual spatial autocorrelation in species distribution models. *Ecography*. 2012; 35(10):879–88. <https://doi.org/10.1111/j.1600-0587.2011.07138.x>
36. Stan Development Team. Stan Modeling Language Users Guide and Reference Manual, Version 2.272021.
37. e-Water Cooperative Research Centre. MUSIC (Model for Urban Stormwater Improvement Conceptualisation) Version 6.0. Canberra, Australia: e-Water Cooperative Research Centre, 2013.
38. Bratières K, Fletcher TD, Deletic A, Zinger Y. Nutrient and sediment removal by stormwater biofilters: a large-scale design optimisation study. *Water Res*. 2008; 42(14):3930–40. <https://doi.org/10.1016/j.watres.2008.06.009> PMID: 18710778
39. Loperfido JV, Noe GB, Jarnagin ST, Hogan DM. Effects of distributed and centralized stormwater best management practices and land cover on urban stream hydrology at the catchment scale. *J Hydrol*. 2014; 519:2584–95. <https://doi.org/10.1016/j.jhydrol.2014.07.007>
40. Government of Victoria. Yarra Ranges Planning Scheme, Schedule 2 to Clause 42.01 Environmental Significance Overlay. Little Stringybark Creek Melbourne, Victoria 2019. https://planning-schemes.delwp.vic.gov.au/schemes/yarraranges/ordinance/42_01s02_yran.pdf.
41. Lake PS, Bond N, Reich P. Linking ecological theory with stream restoration. *Freshwater Biol*. 2007; 52(4):597–615. <https://doi.org/10.1111/j.1365-2427.2006.01709.x>