SQIDEP

Draft in response to initial phase of consultation

Contents

Tables

Figures

Stormwater Quality Improvement Devices Evaluation Protocol (SQIDEP)

Prepared for release on behalf of Stormwater Australia by its Stormwater Quality Improvement Device Advisory Committee (SQIDAC).

Acknowledgements

This document has been developed by Stormwater Australia and is broadly based on the Auckland PDEP.

In preparing this document Stormwater Australia acknowledges the input made by the Auckland PDEP contributors in advancing this important cause.

Stormwater Australia also acknowledges the time and technical contributions made by the members of the Stormwater Quality Improvement Device Advisory Committee (SQIDAC):

- Mr Andrew Allan, Stormwater Australia
- Mr Michael Wicks, Stormwater 360
- Mr Stephen D. Baker, Rocla/Fletcher Building Products
- Mr Andy Hornbuckle, SPEL Environmental
- Mr Andrew Macklin, Ecosol
- Mr Chris Rochfort, Star Water
- Dr Charles Kelly, Humes/Holcim
- Dr Darren Drapper, Drapper Environmental Consultants

1. GLOSSARY

Acronym	Term	Explanation
ADT	Average Daily Trips	Traffic movement count.
AEP	Annual Exceedance Probability	Probability that a given rainfall total accumulated over a given duration will be exceeded in any one year.
	Aliquot	A portion of a larger whole, especially a sample taken for chemical analysis or other treatment. For the purposes of this protocol a discrete sub-sample collected from a qualifying storm event.
APHA	American Public Health Association	Reference organisation.
ARI	Annual Recurrence Interval	Frequency of storm event.
ARQ	Australian Runoff Quality	Document published by Engineers Australia providing guidance on procedures to for the estimation of urban stormwater contaminants and associated design guidelines.
ASTM	American Society for Testing Materials	Reference organisation.
BoE	Body of Evidence	One evaluation route in the SQIDEP, incorporating existing data from other sites.
	Calibration	Utilising monitoring data points to adjust certain parameters used for the sizing methodology to ensure its representativeness.
	Claimant	Designer, vendor or supplier of permanent Stormwater Quality Improvement Device.
CRE	Concentration Removal Efficiency	A mathematical ratio of the difference between an influent concentration and an effluent concentration. Expressed as a percentage. Quantifies the ability of a device to reduce the concentration of a contaminant in stormwater.
	Controlled Field Test	Tests on a full scale device installed in the field, using artificially-produced influent to mimic stormwater flows.
	Effluent (or Outflow)	Stormwater exiting a treatment device.
DQO	Data Quality Objectives	

Device	Stormwater Treatment Device	Any permanent, repeatable man made device, structure or system designed primarily for the improvement of stormwater quality.
DRP	Dissolved Reactive Phosphorus	Any form of P that reacts with reagents in a colorimetric test following filtration of the sample through a 0.45 μm filter paper.
	Evaluation Panel	Independent panel set up to make final decision on whether to certify device performance.
EMC	Event Mean Concentration	Weighted average pollutant concentration that reflects varying runoff concentration over the duration of the hydrograph.
ESA	Equivalent Standard Axles	Traffic movement count.
	Influent (or Inflow)	Stormwater entering a treatment device.
IET/ADP	Inter-event Time (also known as Antecedent Dry Period)	Time between a storm event's end and the subsequent event's beginning as designated by minimum time interval with no greater than 1mm of rainfall.
IQR	Inter Quartile Range	A measure of statistical dispersion, being equal to the difference between the upper and lower quartiles.
	Laboratory Tests – Scale Model	Tests undertaken in the laboratory on a scaled down model of the device. NOTE this data is not accepted for this SQIDEP.
	Laboratory Tests – Full Scale	Tests undertaken in the laboratory on a full scale model of the device.
LPT	Local Pilot Trial	One evaluation route in the SQIDEP, requiring field installation and monitoring of full scale device performance.
MRE	Mass Reduction Efficiency	A mathematical ratio of the difference between the influent pollutant load by mass (i.e. concentration multiplied by flow volume) and the effluent pollutant load. Expressed as a percentage. Allows the total mass of contaminant captured by a device to be quantified.
NATA	National Association of Testing Agencies	Industry peak body responsible for certifying analytical agencies to ensure technical competence in undertake specific testing and analytical methods.
	Performance Metrics	Quantify pollutant removal capacity and consistency of treated effluent water quality.

PSD	Particle Size Distribution	Description of particle sizes (ranges) in stormwater flows.
QAPP	Quality Assurance Project Plan	Plan to show how performance testing in the field is undertaken in a way that ensures appropriate methods and procedures are followed.
RAE	Relative Achievable Efficiency	A performance metric. Expressed as a percentage. Determines pollutant removal relative to an irreducible minimum concentration or a water quality standard.
SFR	Specific Flow Rate	The flow rate through the device divided by the cross sectional area of the device.
SQIDAC	Stormwater Quality Improvement Device Advisory Committee	An advisory committee reporting to the Stormwater Australia board.
SQIDEP	Stormwater Quality Improvement Device Evaluation Protocol	The testing protocol described in this document.
SSC	Suspended Sediment Concentration	A method for measuring sediment in stormwater according to a standard laboratory method (i.e. ASTM D3977-97 Test Method B or equivalent). Note some laboratories may refer to a similar method as a "low concentration TSS". SSC is different to Total Suspended Solids (TSS) and results should not be used/reported interchangeably.
Tc	Time of Concentration	A measure of the response of a catchment to a storm event. It is the longest time required for water to flow from the most hydrologically remote point in a catchment to the catchment outlet. It is a function of the topography, geology, and land use within a catchment.
TKN	Total Kjeldahl Nitrogen	The sum of organic nitrogen, ammonia (NH3), and ammonium (NH4+) in a sample.
TN	Total Nitrogen	The sum total of organic and oxidised nitrogen species (NOX).
TP	Total Phosphorus	Sum of organic and inorganic forms of phosphorus in unfiltered water samples.
TSS	Total Suspended Solids	A method for measuring sediment in stormwater according to a standard laboratory method (e.g. APHA (2005) 2540 D).

		Total Suspended Solids Should not be used/reported interchangeably with SSC. The terminology is distinctly different from Gross Pollutants and should not be misinterpreted as such.
USEPA	United States Environmental Protection Agency	Reference agency.
	Validation	Utilising known data points to confirm a result or prediction.
VPD	Vehicles Per Day	Traffic movement count.
WERF	Water Environment Research Foundation	(United States) Industry body and reference organisation.

2. Introduction

The following process describes the field trial pathway to undertaking testing under Australian conditions.

Subsequent sections describe the requirements for a Quality Assurance Project Plan prior to field testing being undertaken, and the reporting and evaluation of results.

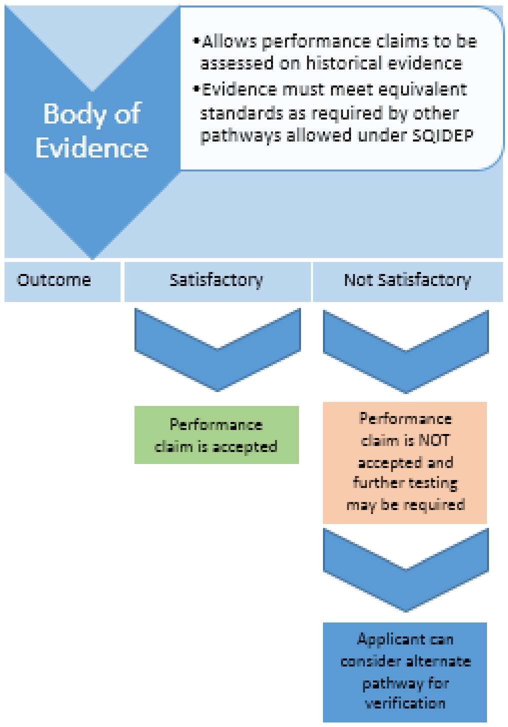
At this stage the SQIDEP does not stipulate who should undertake these evaluations (apart from indicating there needs to be independence from claimant) nor does it describe in detail the requirements for Body of Evidence evaluation. These aspects will need to be addressed in subsequent work aimed at producing further guidance.

The primary focus of this protocol is to allow field testing and subsequent evaluation of results, and has been set up to provide a framework to support activity into the future. It is recognised that in many instances there exists a body of prior research and development which would be a candidate for a Body of Evidence evaluation; it is not the intention of the SQIDEP to dismiss this previous information, however in order for it to be considered it should demonstrated to have been done to an equivalent standard as outlined in this protocol.

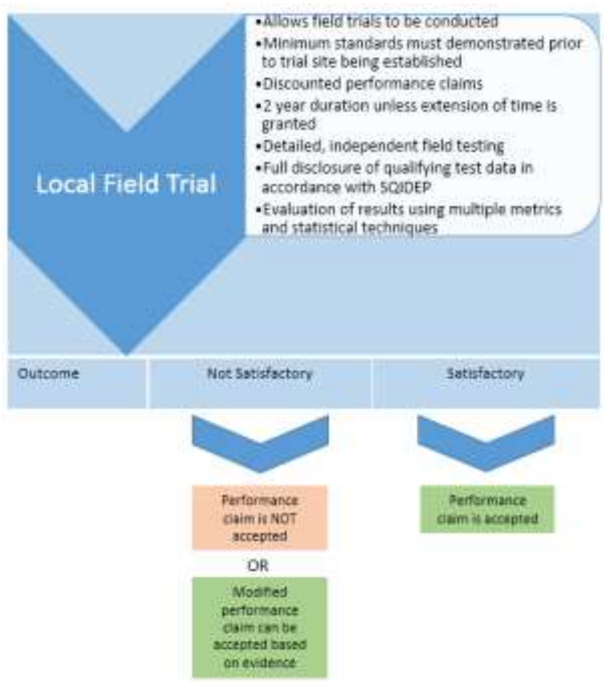
3. Process overview

The following diagram show the process for performance results to be considered under either a Body of Evidence or Field Evaluation pathway.

This protocol deals primarily with aspects of Field Evaluation.



Body of Evidence Pathway



Field Evaluation pathway

Figure 1. Evidence Pathways

4. Prior to commencing field testing

Prior to undertaking a detailed set of field evaluation it is important that there is sound evidence available to support the design of the trial. This would typically include laboratory testing (especially using a validation framework which included a range of full scale and challenges tests) or a solid theoretical base which outlined key treatment mechanisms and credible scale/ pilot laboratory tests.

It is up to the claimant to identify a suitable trial site and convince the site operator of the veracity of the claims in these negotiations. Under this protocol the claimant can be afforded a reduced performance claim (as a precaution against ambit claims to enter and become established in the market) and two years to collect data and undertake analysis to prove the claim (unless an extension of time is justified), however the design of the field testing regime should be based on the full treatment claim to ensure sizing relationships (i.e. between catchment and devices) are commensurate with expected recommendations should the claim be proved.

If the claimant is able to demonstrate the theoretical, laboratory or field performance of the device, based on a desktop assessment, then a Quality Assurance Performance Plan should be developed to support more rigorous field trials.

5. Quality Assurance Project Plan

A Quality Assurance Project Plan (QAPP) must be prepared prior to field testing. Its purpose is to describe how performance testing is to be conducted and has a key objective to ensure appropriate methods and procedures are followed and documented so that data obtained during testing is valid for verification of the device performance.

The QAPP must be prepared by the claimant and agreed by the independent evaluator before testing commences. The plan may be revised as necessary throughout the course of field testing with adjustments, notes and explanation provided with the consent of the independent evaluator.

The QAPP should contain background information on the device being tested, project organisation, sampling design and methods, laboratory methods, field and laboratory quality control, data management procedures, data review, and reporting.

The QAPP is developed for the purpose of planning the monitoring programme and to ensure that the proposed methodologies are executed in line with the contents of the QAPP, which are aligned with the protocols outlined in the SQIDEP.

The QAPP must be agreed by the claimant who should also commit adequate resources to implement the recommended testing. The QAPP should be developed by a person with knowledge of the SQIDEP and a good understanding of field sampling and analytical chemistry methods. Where appropriate, it shall be developed in consultation with the analytical laboratories selected, especially if specialist analysis is required.

The QAPP shall be based on the claimant's Performance Claim, and shall contain the details of:

a. Data Quality Objectives.
b. Organisational roles and responsibilities.
c. Description of test site.
d. Measuring rainfall.
e. Storm events sampled.
f. Flow monitoring.
g. Sampling location.
h. Sampling equipment.
i. Sampling methodology.
j. Sampling Quality Assurance and Quality Control.
k. Laboratory analysis.
l. Laboratory Quality Assurance and Quality Control.
m. Data management.
n. Reporting.

For each of these items (a- n) the consultation SQIDEP document provides further commentary, which should be retained. Where the commentary contains numerical requirements (e.g. numbers of samples) these should be updated to reflect the numbers presented later in this document Table 7.

The QAPP shall describe the procedures that will be used to ensure data quality and integrity. The QAPP shall detail how the following will be achieved in accordance with recognised publications which are equivalent to, or complement accepted EPA methods.

While the primary focus of the QAPP is to ensure collection of relevant, quality data for the purposes of evaluating performance claims, it is the responsibility of all parties involved to ensure that all activities are undertaken in a manner consistent with occupational health and safety considerations.

5.1. Data Quality Objectives

The Data Quality Objectives (DQO) are to obtain accurate and relevant data to assess the claimant Performance Claim. Data quality will be assessed against the criteria of representativeness, completeness, and comparability.

Where a device provides quantity control, it is essential that monitoring at a site considers both contaminant concentration and mass or load transported. Data collected must be representative of the majority of storm event sampled (and in particular, rising and falling limbs of the storm hydrograph), for each event that forms part of the device evaluation. Representativeness is largely achieved through the collection of flow-weighted event mean concentration samples, except for those contaminants that are obtained by grab samples.

The events sampled must also represent rainfall and thus runoff patterns for the catchment across an extended period of time (> 12 months) and be subject to the qualifying number of characteristic storms being achieved. Representativeness shall be assessed and reported.

Completeness of data will require that enough storm events are sampled to allow accurate evaluation.

The data collected must be able to be compared to performance at other locations. Comparability requires that the contaminants analysed (e.g. total suspended solids versus suspended solids concentration) and the sampling and analytical methods used can be compared.

5.2. Organisational Roles and Responsibilities

There are many different parties involved in measuring device performance. These generally include the regulatory body, claimant, general contractors (including installation and maintenance contractors), testing organisation, analytical laboratory, site owner, and the evaluation panel. All have roles and responsibilities in the successful completion of a project (USEPA, 2002).

Organisational roles and responsibilities shall be clearly identified in the QAPP. The claimant, sampling organisation (including both equipment and sampling), analytical laboratory, and reporting organisation shall be clearly identified, along with limits of their roles. Ideally, key personnel and their titles and contact information will be included. An organisational chart should clarify personnel and their roles (especially in confirming independence requirements).

5.3. Description of Test Sites

Ideally, a test site shall be selected so the results can be applied elsewhere. The claimant shall propose a suitable site and demonstrate its appropriateness for performance testing to the evaluator.

This site shall be representative of the installation and land uses appropriate to the device and intended market segments. The test site land use shall be detailed and described according to land use category).

It is considered good practice that the trial site water quality is examined to provide confidence that runoff quality will be appropriate to allow treatment efficacy to be demonstrated. Catchments delivering 'clean' water near or below irreducible concentrations or analytical method detection limits are unlikely to provide sufficient numbers of qualifying samples to allow definitive assessments to be undertaken.

Guidance on expected catchment characteristics should be sought from publications such as

Within datasets being presented for analysis there should be limits on maximum and minimum pollutant concentrations that can be included within the set of qualifying storms. Where a device has demonstrated its ability to achieve pollutant removal at low concentrations (i.e. below those given in Table 5-1), these results can be reported but should not be included in determining if a storm is a qualifying events. To avoid committing to testing at a site where pollutant concentrations are likely to fall outside the minimum and maximums allowed for qualifying events or where pollutant concentrations fall below laboratory Limits of Detection, it is recommended that the claimant take samples to characterise the trial site. As an indication, stormwater should contain the average typical concentrations of contaminants as provided in recognised industry publications and studies (i.e. such as Australian Runoff Quality). Typical lower concentrations for commonly regulated pollutants are provided in **Table 6-2**[1].

It is acknowledged, however, that pilot sites may be limited to new developments through the regulators development assessment process. Therefore, testing should continue for a period sufficient to demonstrate a range of typical catchment pollutant concentrations, and when this data is collected it can be used to augment concentration referenced in other literature or to develop a site specific dataset (i.e. in the case that there is a paucity of published data for the particular application).

[1] It should be noted that the SQIDEP is intended to provide a framework that is adaptable to allow the assessment of other pollutants

Table 5-1: Typical Untreated Stormwater Contaminant Concentrations for Selecting SQID Trial sites

	Recommended minimum influent concentration (C*) mg/L[1]	Recommended Mean Influent Concentrations & (Standard Deviation)[2]	Mean + 1SD	Mean + 2SD
TSS	6	151 (+220)	371	591
TP	0.06	0.34 (+0.37)	0.71	1.08
TN	1.0	1.82 (+1.27)	3.09	4.36

Notes:

1. Recommended minimum influent concentrations or C* Values Based on all Parameters in the 2008 International BMP Database Summary (Geosyntec and Wright Water Engineers, 2008, eWater, 2010)
2. Recommended mean influent concentrations and standard deviations from Goonetilleke, A, Thomas, E, Ginn, S, and Gilbert D (2005), Understanding the Role of Land Use in Urban Stormwater Quality Management.

The recommended mean influent concentrations are given as guide and may lead to disqualification of a pollutant parameter for an individual storm if agreed between the technical expert reviewer and independent evaluator. The mean concentration for both Nitrogen and Phosphorus for storm events collectively must be less than the mean influent concentrations + 1 standard deviation as shown above. Up to 20% of individual storms must be less than the mean influent concentrations + 2 standard deviations as shown above. Any mean influent concentrations above these values will be automatically disqualified from the analysis.

A full description of the test site shall be provided, and shall include the following:

a. Catchment area, land use, percentage impervious cover;

b. Aerial photos and site photos;

c. Geology, hydrogeology, soil types, surface hydraulics;

d. Potential pollutant sources;

e. Baseline stormwater quality;

f. Site map, showing catchment area, drainage system layout, treatment device, and sampling points, preferably GIS compatible;

g. Treatable flow rate;

h. Expected catchment flows;

i. Make, model and capacity of treatment device;

j. Closest receiving water body;

k. Identification of bypass flow rates and/or flow splitter design;

l. Pre-treatment system, if any;

m. Site suitability – e.g. safety, access for flow measurements, power, phone; and

n. Any known adverse site conditions.

5.4. Measuring rainfall

Rainfall shall be measured by a rain gauge (pluviometer) that is capable of sampling at intervals of 5 minutes or less, and in increments no greater than 0.25mm. An electronic rain gauge connected to a data-logger (pluviometer) is recommended. A non-recording rain gauge installed at the test site will allow the recording gauges totals to be calibrated and increase confidence in data.

The location of the rain gauge in relation to the test site shall be shown on a map. The rain gauge shall be installed and maintained according to manufacturer's instructions, and as a minimum be checked and cleared of debris after each sampling event and calibrated at least two times during the testing period (if a non-recording gauge is used this can be emptied and 'reset' to achieve this). It is also recommended that rain gauges are checked prior to each anticipated storm event targeted for sampling to ensure they have not clogged and rainfall data is recorded during the monitoring event. Rain gauges shall be protected from excessive wind velocities that could skew accuracy of measurement.

Guidance on installation of rain gauges can be found at the Bureau of Meteorology website http://www.bom.gov.au/climate/how/observations/rain-measure.shtml.

5.5. Qualifying Storm Events

A sufficient number of qualifying storm events required for a statistically robust data set to support analysis. In addition to achieving a sufficient number of events, the data set should ensure a range of flow conditions are demonstrated.

Pollutants that are generated with a level of consistency may require a smaller number of events to be analysis to achieve statistical confidence for performance. Where pollutant generation is more variable additional samples may be required to account for this variability in influent concentration.

In all cases a minimum of 8 qualifying events is required, but an upper number of tests needs to be determined based on an assessment of the data using credible statistical methods (such as ANOVA/ t-test techniques) to achieve at least 90% statistical significance between paired samples of influent and effluent. If the level of statistical significance is not able to be demonstrated more events must be sampled until the 90% statistical significance is achieved. Where this statistical validation may require excessive events (e.g. >30) to prove (as determined by the recognised statistical power techniques such as the equation described by Burton and Pitt, 2001) an altered claim can be considered commensurate with evidence.

At least three (3) of peak runoff flows from the sampled events should exceed 75% of the design flow rate of the device and one at or greater than its design flow rate. Where a site is expected to generate pollutants in concentrations that are demonstrably different from other land use categories (e.g. pollution associated with specific types of industrial runoff) and a treatment effect is being claimed, the design of the field tests should account for this fact.

Claimants may alternatively choose, with the approval of the evaluator, to conduct the field trials on multiple sites, with a statistically significant number of sampling events on each test site. Subject to being able to demonstrate statistical comparability between multiple sites (especially in relation to inflow) results could be ‘pooled’ to increase the power of statistical techniques. A minimum of 90% statistical significance between influent and effluent treatment reductions is the goal. Ideally this would allow results collected across multiple sites to augment the knowledge base where it is demonstrated that key environmental parameters are within similar ranges.

Qualifying storm criteria are defined for three features, the minimum storm depth, Antecedent Dry Period (ADP) and minimum runoff duration.

An ADP of three days (72 hours) of no rainfall is recommended to differentiate between the end of one storm and the beginning of another. In making this recommendation 72 hours has been chosen to achieve a balance between catchment pollutant generation conditions and the practicalities of collecting and submitting samples for analysis, however shorter ADPs are still able be considered if qualifying event criteria are met.

Sampling events should be distributed throughout the monitoring period sufficiently to capture seasonal influences on storm conditions and device performance.

There is no stipulated minimum storm event duration, however for the majority of qualifying events (80%) at least 8 aliquots is required. Aliquots shall be collected across the storm event to ensure samples are representative of both rising and falling limbs of the hydrograph. It is expected that limitations of the sample collection equipment (i.e. time to fill sample containers, time to reset between samples) will in part dictate a minimum storm duration. Fewer aliquots may be acceptable (min. 3) on small rainfall and runoff events (eg. <8mm) provided they are sampled from the rising limb, peak and falling limbs of the hydrograph (ie. not biased to one limb only). Overall, the full set of storms should contain longer and shorter duration storms covering different flow scenarios.

The QAPP should allow results to be presented and interpreted along with qualifying storm characteristics as summarised in **Table 7**.

5.6. Flow Monitoring

Flow monitoring equipment must be able to continuously monitor flow at regular intervals to match rainfall information (5 minutes or less is recommended) throughout the duration of a storm event and over the expected range of flows. Depth measurement or area/velocity devices are most common – selection will depend upon the test site and method of conveying stormwater (WSDE, 2004).

The flow at the inlet and outlet shall be measured. Monitoring of bypass flows is optional, however at a minimum the monitoring information should be sufficient to identify periods during which the device is operating in bypass mode. Proper presentation and interpretation of passing and bypassing flows is considered critical to ultimately supporting design outcomes.

Flow monitoring equipment shall be described (make and model), and the description shall include flow splitter and bypass flow set points, and flow conditions (gravity or pressure) (WSDE, 2004). Equipment must be installed and calibrated according to manufacturer's instructions and as described in the QAPP.

As a minimum, the equipment's internal desiccators, sensors and connections shall be inspected between each sampling event (USEPA, 2002). It is also recommended that in-pipe sensors and intakes be checked prior to each anticipated storm event targeted for sampling to ensure they are not blocked, damaged, or covered by sediments or gross pollutants.

When determining suitable flow monitoring configurations the relevant Australian Standards should be consulted and references provided.

5.6.1. Accounting for internal bypass flows

Some devices by nature of their internal geometry and treatment process may allow a portion of flows entering to be internally bypassed. In these instances flow monitoring should be sufficient to determine the treatment effectiveness across the devices (i.e. inlet and outlet conditions).

5.7. Sample Location

The inlet sample shall be taken as close as possible to the device, at a minimum at a point where total site runoff is sampled. Likewise, effluent characterisation should account for total storm flow, including bypass if it occurs. For basic and enhanced treatments, gross pollutants (>500 µm) should be excluded (WSDE, 2004) from any captured water samples, unless this is being claimed for the device.

Outlet flow should be sampled either prior to or after mixing with bypass flow (Figures 5-1 and 5-3.

In either case, the contribution of bypass (if/when it occurs) shall be incorporated into the calculation of device efficiency (USEPA 2002) or design tools as appropriate as described below.

If the outlet flow is sampled prior to mixing with bypass flow (Figure 5-2a) it should be noted when the bypass condition occurs (but it is not necessary to measure bypass flows). The performance claim must be made in relation to the device, and no removal can be claimed for the bypass flows. In this circumstance the performance claim claimed must be qualified as such; sizing and design advice must recognise this fact.

If the outlet and bypass flow is to be sampled together (Figure 5-3b), samples should be collected after sufficient mixing has occurred and prior to comingling with any other runoff. In this event bypass flows must be measured and the concentration of the bypass flows assumed to be the same as the device influent.

Figure 5-3 shows sampling and flow monitoring configurations for devices that allow internal bypass of entering flows. Often this may be due to changes in treatment veracity over the device's deign life (e.g. filters becoming clogged reducing hydraulic throughput).

If internal bypass occurs and samples are collected immediately after the treatment element (i.e. separate to internal bypass flows) additional flow monitoring will be required to allow the treatment effect for all flows entering the device to be calculated.

Figure 5-1. Flow Sensor and Sample Intake Locations (bypass flows not accounted for in analysis)

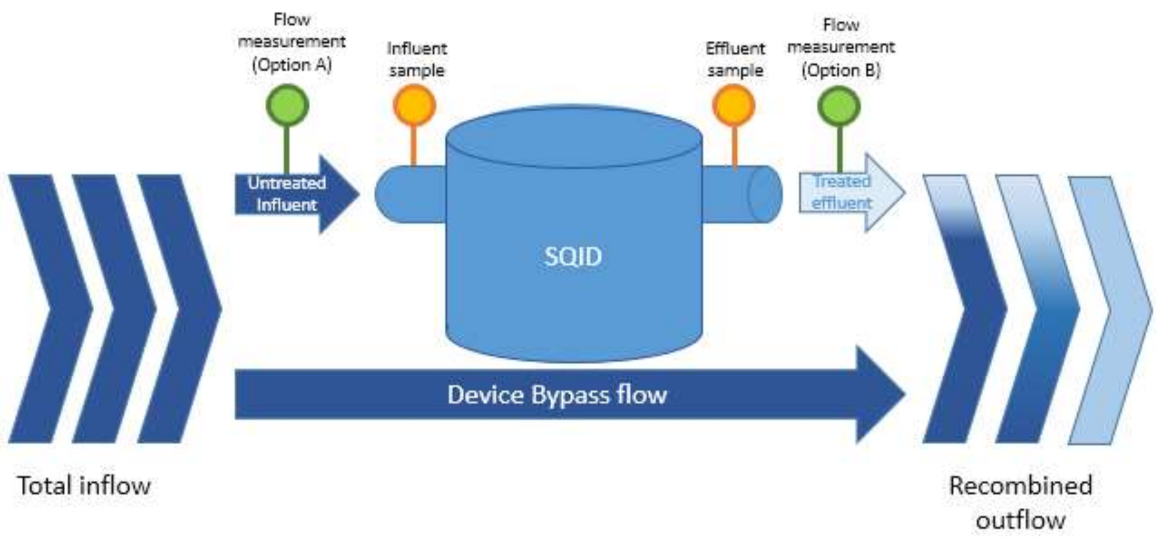


Figure 5-2a. Flow Sensor and Sample Intake Locations (bypass flows accounted for in analysis)

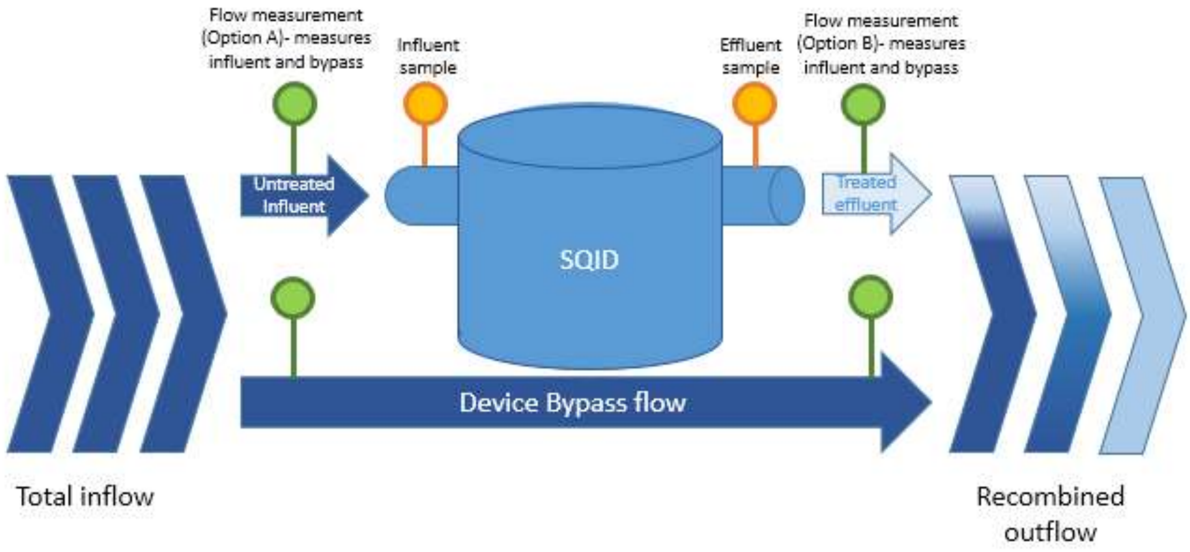


Figure 5-2b. Flow Sensor and Sample Intake Locations (bypass flows accounted for in analysis)

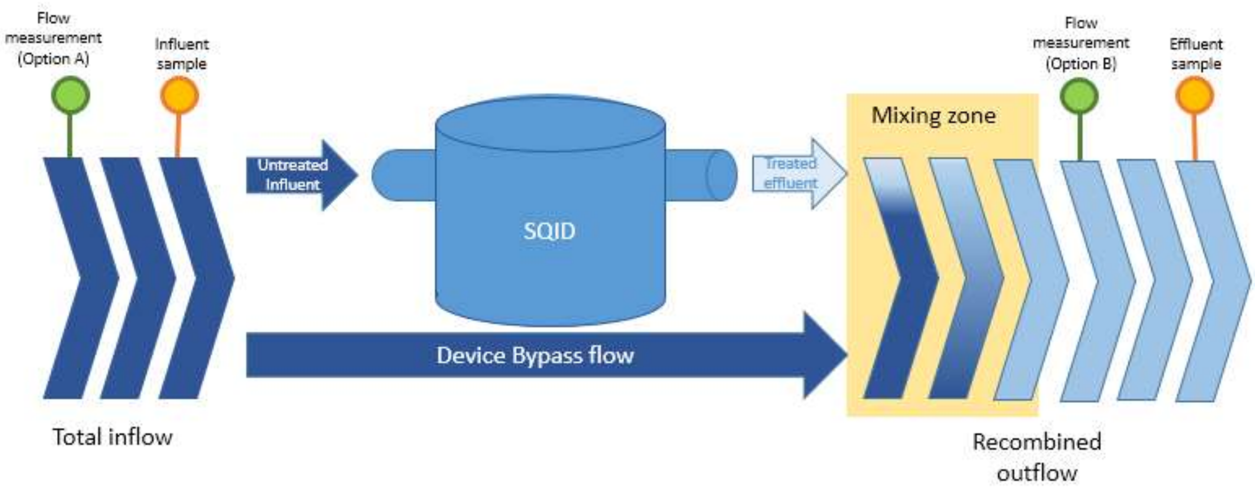
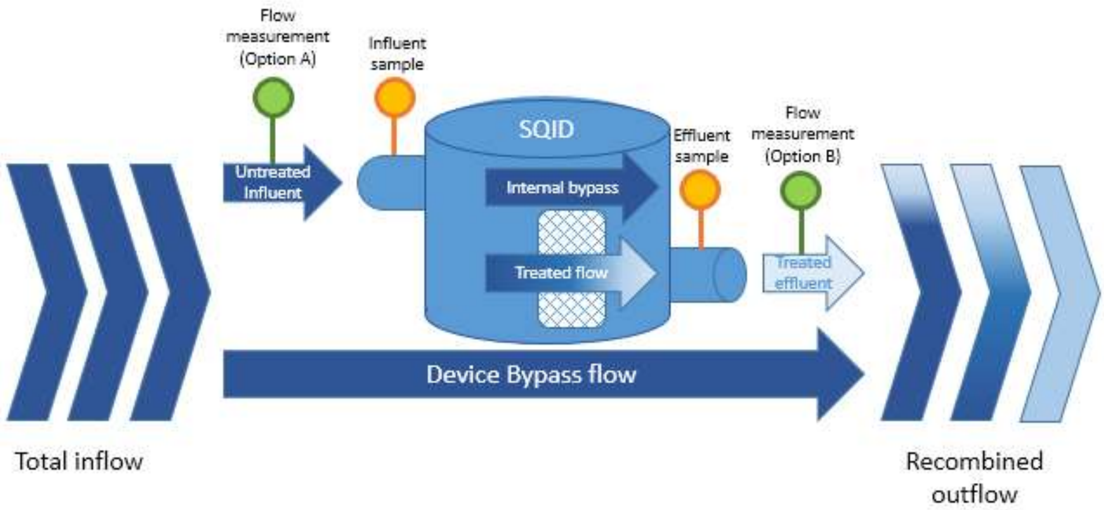


Figure 5-3 Sample location accounting for any internally bypassing flows



There is potential for some stormwater constituents to stratify during conveyance. To avoid sampling stratified flow, all sampling points shall be located where mixing of the flow is maximised (USEPA, 2002; WSDE, 2004). Sampling locations should be consistently located above and downstream of the tested device to allow representative consideration of stratification. It is recommended the location is agreed through the QAPP to accommodate operational realities of field testing.

Hydrocarbons or other light non-aqueous phase liquid which are likely to remain in a floating, free state at time of arrival at the testing site shall be sampled in accordance with recognised guidelines. Where emulsified hydrocarbons are expected, justification should be provided and samples should be collected from a zone of representative mixing with appropriate collection and preservation techniques used.

Any variations to sample location must be approved in the QAPP.

5.8. Accounting for scour

For devices installed online, scouring might occur during large events. Any scouring effects shall be assessed and reported. The assessment may be in the form of hydrodynamic modelling. The claimant may alternatively propose a different methodology of assessing scour in consultation with the evaluator. Otherwise if the device is an offline device; and/

or there is sufficient evidence that scouring is not present, it shall be provided as part of the Performance Report.

In situations where there is the potential for scour to occur preventative strategies can be recommended as part of installation or maintenance methodologies.

5.9. Sampling Equipment

Evaluation of device performance requires measurement of stormwater inflow into the device, outflow, stormwater quality, and rainfall.

Equipment is required to measure rainfall, inflow and outflow volumes, and some method of determining the bypass volumes must be incorporated (measurement or calculation). Equipment is also required to sample stormwater for laboratory analysis. For all equipment, the claimant and model of equipment, and procedures and schedule for calibration, inspection and cleaning shall be provided (USEPA 2002).

Consideration should be given to access for sampling equipment, equipment security and protection, and power (access to grid, or unobscured sky for solar) and phone supply (if the site is to be remotely telemetered, either land line access or cellular reception). The potential for power failure and subsequent loss of samples should also be considered (WSDE, 2004).

5.9.1. Automated Samplers

Automated samplers are to be used for all sampling, except where grab samples are required (i.e. to ensure timely sample preparation, preservation or monitor unstable parameters).

The sampler shall be installed and calibrated according to the manufacturer's instructions, and maintained between each sampling event. Information provided shall also describe how the sampler will be programmed, how sampling will be triggered, and how the sampler purges and rinses between samples (USEPA, 2002; WSDE, 2004).

The bottle changing procedure shall be described. The suction tube material, length and vertical lift should be described, and the location of the tube inlet relative to flow conditions should also be described. Teflon shall be used for sampling organic constituents (WSDE, 2004).

5.10. Sampling Methodology

A flow weighted composite sample should be collected utilising an automated sampler whenever possible. However, some contaminants may require grab sampling depending on the preservation methods required under sampling protocols.

5.10.1. Grab Sampling

Grab stormwater samples are discrete samples (not composited), normally collected within the first 30 minutes after the onset of runoff, but no later than within the first 60 minutes. If grab sampling is required, the approach shall be clearly documented. The QAPP shall describe how the criteria for a qualifying storm event shall be met. The availability and preparedness of sampling staff shall also be demonstrated.

Grab sampling is required for constituents that transform rapidly, require special preservation, adhere to bottles, or where compositing can mask the presence of some contaminants through dilution. Grab sampling is required for pH, temperature, total petroleum hydrocarbons (TPH), oil and grease, mercury (Hg), hexavalent chromium (Cr^{6+}), bacteria, cyanides, total phenols and residual chlorine. For all other constituents and pollutants, sampling shall be undertaken using automated samplers. Samples to be analysed for TPH shall be collected directly into the bottle that will be used in the laboratory. Samples for bacteriological analyses shall be collected into sterile bottles using appropriate clean sampling and equipment handling techniques in line with relevant EPA guidelines. Any deviation to these sampling requirements must be approved in the QAPP.

5.10.2. Automated Sampling

Where the constituent being measured does not require grab sampling, automated sampling should be undertaken. Samples can be taken by automatic flow-weighted compositing, or discrete samples that are composited later. Where samples are manually composited, it is recommended this is undertaken at the analytical laboratory to minimise risks of contamination.

5.10.3. Time-Proportional Sampling

Examples of good and poor hydrograph representation for time-proportional sampling are attached in the Appendices. This is a method of sampling that is best suited to auto sampling techniques.

A statistically representative number of discrete samples or sample aliquots shall contribute to each composite sample, with the emphasis on the hydrograph's rising limb (Fassman, 2010), at both the inlet and outlet of the device.

The sampler should be programmed to take the maximum number of aliquots possible (USEPA, 2002). All samples collected from qualifying events should be analysed and reported.

While time-proportional sampling is acceptable, flow proportional compositing of the samples must be undertaken. The flow data will need to be readily available following a storm event to properly composite time-proportional samples and ensure holding times are met.

5.10.4. Flow-Proportional Sampling

As many aliquots as possible should contribute to the composite sample (Fassman, 2010; Ma et al, 2009) and should provide statistical confidence in representativeness.

5.11. Sampling Quality Assurance and Quality Control

Operation and maintenance schedules for sampling equipment (e.g. automated), flow monitoring and rainfall equipment shall be provided.

Sample blanks for field and analytical testing will be supplied in accordance with QAPP and in line with recommendation in EPA guidelines.

Chain of custody documents identifying sample, collection agency, collection time, preservation used and laboratory receipt of sample and sample condition shall be provided.

5.12. Laboratory Analysis

All analysis shall be undertaken at laboratory or analytical facility with current NATA accreditation for requested analysis (including limit of reporting).

The method chosen for analysis shall be determined in the QAPP, including any justifications as considered necessary (e.g. depending on expected catchment conditions, analysis methods maybe chosen on the basis of limit of reporting).

5.13. Laboratory Quality Assurance and Quality Control

Proper quality assurance and control procedures are critical within any laboratory engaged to undertaken samples.

Generally, the use of NATA accredited facilities will ensure that a high standard of quality management is adhered to, but beyond the accreditation for specific tests the laboratory should be able to provide suitable chain of custody documentation to identify sample receipt and condition, samples should be properly labelled and stored pending testing, and holding times for samples should be observed.

In addition to field based quality assurance samples, the laboratory should have its own procedures to demonstrate confidence in sample preparation methods and analysis including the use of duplicates, spikes, surrogates and blanks.

When examined together, quality assured laboratory and field data give the highest confidence in the measured results.

5.14. Data Management

All documentation pertinent to the undertaking of the field testing, sample collection and analysis and reporting of results should be retained in full and presented in a logical and easy to follow format for evaluation.

It is desirable to receive testing results in electronic format to facilitate analysis and assessment, but copies of the accompanying certificates of analysis which include test results and laboratory quality assurance results should also be retained, along with Chain of Custody documentation and any relevant field notes identify sample collection time, location and prevailing conditions.

Where electronic copies are provided these should be delivered to the independent party who is working with the claimant to deliver field testing.

Full copies of original results (electronic and hardcopy) should be retained in order to be made available to any later evaluating party.

5.15. Reporting

Reporting must be undertaken against the approved QAPP by an organisation independent of the claimant. The reporting organisation must have an understanding of the hydraulics and treatment mechanisms of the SQID, with knowledge and experience of proper sampling and flow measurement practices, and with the ability to properly interpret and report, without prejudice, the flow and water quality data.

A statement verifying the independence of the report author and any agency involved in testing must be supplied.

A sample template for preparing a Performance report will be provided as an Appendix.

6. Performance Reporting

The performance of a device needs to be reported consistently for efficient and accurate evaluation. This section discusses the requirements of the SQIDEP in terms of the framework of the report and the performance metrics.

6.1. Non-Detects

Non-detects are values reported to be at or below a reporting limit and/or detection limit. They need to be considered when analysing data as removing them may result in biased and non-representative estimated summary statistics of the monitored site (Helsel, 1988).

In the SQIDEP, non-detects are reported against the detection limit in accordance with internationally accepted methods. If there are a large number of non-detects, the applicant can propose the use of a more reliable statistical method to analyse them, such as (1) regression on order statistics (ROS) method; (2) maximum likelihood estimation (MLE) method (Kayhanian, 2011).

6.2. Framework for Reporting

Performance reporting is required after the LPT is completed. Devices evaluated using the BoE route are also required to summarise existing data and report using the following framework.

The requirements for reporting are as follows:

a. Device information (extracted and summarised from AP Report);
b. Sizing methodology and its description, including any non-validated or non-referenced assumptions;
c. Sampling and analytical methodologies (extracted and summarised from QAPP);
d. Data reporting (for all qualifying events);
e. Discussion of any factors affecting the performance, including scaling effects and particle size distribution of both the influent and effluent. Other factors shall be included if deemed appropriate;
f. Box and Whiskers Plot for the Influent and Effluent Flows;
g. Statistical Significance Testing; and
h. Analysis of Non Detects if applicable. Conclusions and Recommendation;
i. Data quality (below); and
j. Performance metrics (below), results and discussion.

6.3. Data Quality

The data collected shall be assessed and reported for the following factors:

a. Representativeness, completeness and applicability of rainfall/ runoff; and
b. Values relative to the detection limits of the analytical methods applied.

6.4. Performance Metrics

The pollutant removal capacity of a device needs to be consistent, and provided suitable information is collected at the time of field trials, multiple metrics are able to be determined and should point to a consistent interpretation for the highest levels of confidence in evaluating results.

The SQIDEP allows a number of performance metrics to be presented as follows:

a. Five (5) types of percent removal efficiencies;
b. Event Mean Concentration (EMC) and (if applicable) Mass Discharge Variability; and

c. Statistical significance of differences (if any) between inlet and outlet EMCs.

The details of each performance metric are outlined in Section 14.24 below.

6.4.1. Performance reliability and the statistical analysis of data

A qualifying of number of sampling events is required to verify the statistical representativeness of the removal efficiency. All performance metrics are supported by analysis of data collected when following this protocol, and all should provide a supporting case for the final accepted removal efficiency.

Performance reliability can be measured statistically by several methods. It is assumed that the pollutant concentrations are likely to be log-normally distributed, however this assumption should be verified through statistical techniques and appropriate techniques employed to prepare the dataset for analysis.

Statistical parameters for evaluating the performance of the device seek to understand the difference between paired influent and treated effluent samples (i.e. treatment effect).

a. Ensure that the 90% Confidence Interval of the arithmetic average is provided (CRE and/or MRE calculated as recommended). A Confidence Interval of greater than 90% is required for a claim to be considered valid.

Devices which are able to demonstrate reduced levels of variance in the treated effluent are likely to be able to perform more predictably. Standard statistical techniques can be used to estimate the variability in a dataset. One such procedure is provided below.

b. Measure the spread of the effluent data by analysing the distance of the lowest and upper most point from the 1st and 3rd quartile values (effluent EMCs) against the inter-quartile range (IQR). Within 1.5 times IQR is desired.

c. Calculate the arithmetic mean above and below the standard deviation (CRE and/or MRE). Within one standard deviation is desirable.

d. Calculate the difference between the arithmetic average and the median (CRE and/or MRE). Within 10% is desired.

6.4.2. Removal and Performance Metrics

The methods considered under the SQIDEP to compute and analyse removal rates and efficiencies are:

- Concentration Removal Efficiency (CRE);
- Mass Reduction Efficiency (MRE);
- Relative Achievable Efficiency (RAE);
- Summation of Loads;
- Efficiency Ratio (ER); and
- Flow Based Variability (FBV) Curve;
- Event Mean Concentration and Mass Discharge Variability

Analysis should clearly indicate how treatment and bypass flows (either external or internal to the device) have been accounted for in the presentation of results.

6.4.3. Average and Median Concentration Removal Efficiency

Pollutant Concentration Removal Efficiency (CRE) is computed to determine the reduction in pollutant concentration in a device. Calculations depend on the sampling equipment configuration, as per Figures 14.1- 14.2. The formula for computing CRE is as follows:

Equation 1:

$$CRE(\%) = \frac{EMC_{in} - EMC_{out}}{EMC_{in}} \times 100$$

Where:

EMC_{in}	is the event mean concentration measured in the inflow for each event; and
EMC_{out}	is the event mean concentration measured in the corresponding total outflow for each event. For Sampling Configuration shown in Figure 14.1, EMC_{out} is the event mean concentration measured in the treated effluent. Note, under this interpretation, if bypass is not measured, no credit can be reliably claimed for bypass and design guidance should allow bypass to be excluded from treatment. If bypass occurs and is measured: For Sampling Configuration shown in Figure 14.2a, EMC_{out} should be calculated using Equation 2. For the Sampling Configuration shown in Figure 14.2b, EMC_{out} will automatically include any bypass flow if/when it occurs and Equation 1 can be used.

If an alternative when bypass occurs and it is measured, EMC_{out} for the event is calculated as:

Equation 2:

$$EMC_{out} = EMC_{treatedout}\left(\frac{V_{treatedout}}{V_{totaloutflow}}\right) + EMC_{in}\left(\frac{V_{bypass}}{V_{totaloutflow}}\right)$$

Where:

$EMC_{treatedout}$	is the event mean concentration measured in the treated effluent for the event
$V_{treated\ out}$	is the measured flow volume treated by the device (not bypassing)
V_{bypass}	is the flow volume of the event that bypasses the treatment device (measured or calculated)
$V_{total\ outflow}$	is the flow volume downstream of the junction of the bypass and treated effluent, as per Figure 14-2a, b (measured and/or calculated).

To calculate and report the Average CRE:

1. Calculate CRE for each event, according to Equation 1.
2. Calculate the arithmetic average of the CRE over all events.
3. Calculate the 90% confidence interval for the arithmetic average of CRE.

To calculate and report the Median CRE:

4. Calculate the median CRE over all events.
5. Compute the difference between the arithmetic average CRE and the median CRE.
6. Calculate the arithmetic mean above and below the standard deviation for CRE.

Close agreement of median and average CRE indicate that the overall statistic is not influenced by an extreme event/s. If median and average values are greater than 10% different, the data set should be inspected for the presence of an extreme value(s) which may need further investigation or explanation.

6.4.4. *Average and Median Mass Removal Efficiency*

Pollutant Mass Reduction Efficiency (MRE) is reported to determine the total mass of pollutant captured by the device. MRE calculations are relevant for devices which may provide runoff quantity management. The formula for computing MRE is as follows:

Equation 3: MRE

$$MRE(\%) = \frac{(\mathrm{V}_{in} \times \mathrm{EMC}_{in}) - (\mathrm{V}_{out} \times \mathrm{EMC}_{out})}{(\mathrm{V}_{in} \times \mathrm{EMC}_{in})} \times 100$$

Where:

V_{in}	is the flow volume of each event, measured at the inlet;
V_{out}	is the total outflow volume of each event, measured downstream of the junction of the bypass and treated effluent, as per Figure 14.2
EMC_{in}	is the event mean concentration measured in the inflow for each event; and
EMC_{out}	is the event mean concentration measured in the total outflow for each event, as described by Equation 1 or Equation 2.

To calculate and report the Average MRE:

1. Calculate MRE for each event, according to Equation 3.
2. Calculate the arithmetic average of the MRE over all events.
3. Calculate the 90% confidence interval for the arithmetic average of MRE.

Calculate the median MRE over all events.

4. Compute the difference between the arithmetic average MRE and the median MRE.
5. Calculate the arithmetic mean above and below the standard deviation for MRE.

Close agreement of median and average MRE indicate that the overall statistic is not influenced by an extreme event/s. If median and average values are greater than 10% different, the data set should be inspected for the presence of an extreme value(s) which may need further investigation or explanation.

6.4.5. *Average and Median Relative Achievable Efficiency*

The relative achievable efficiency (RAE) is computed to mitigate the influence of influent EMC on the percent removal calculations. The RAE is a function of benchmark or 'irreducible' concentration. This was derived from the 'best' median effluent concentration across all stormwater treatment devices reported in the International Stormwater BMP Database (Geosyntec and Wright Water Engineers, 2008). The RAE is calculated as follows:

Equation 4: RAE

$$RAE(\%) = \frac{EMC_{in} - EMC_{out}}{(EMC_{in} - C^*)} \times 100$$

Where:

C*	is an irreducible concentration used as a benchmark and taken from Table 14-3 (Fassman, 2010);
V_{out}	is the flow volume of each event, measured at the outlet
EMC_{in}	is the event mean concentration measured in the inflow for each event; and
EMC_{out}	is the event mean concentration measured in the total outflow for each event, as described by Equation 1 or Equation 2.

To calculate and report the Average RAE:

1. Calculate RAE for each event, according to Equation 4.
2. Calculate the arithmetic average of the RAE over all events.
3. Calculate the 90% confidence interval for the arithmetic average of RAE.

To calculate and report the Average RAE:

4. Calculate the median RAE over all events.
5. Compute the difference between the arithmetic average RAE and the median RAE.

Table 6.2: Recommended C* Values Based on all Parameters in the 2008 International BMP Database Summary (Geosyntec and Wright Water Engineers, 2008, eWater, 2010)

Parameter	C*	Based on Treatment Device(s)	Number of Studies
Total Suspended Solids	6 mg/L	Media filter	33
Total Dissolved Solids	N/A[2]		
Total Phosphorus	0.06 mg/L as P	Media filter	28
Dissolved Phosphorus	0.05 mg/L as P	Retention pond Constructed wetland	12 4
Total Nitrogen	1 mg/L as N	Biofilter	12
Total Kjeldahl Nitrogen	0.77 mg/L as N	Biofilter	22
Total Nitrate	0.20 mg/L as N 0.25 mg/L as N	Wetland basin Retention pond Wetland channel	5 12 3
Total Nitrate + Nitrite	0.05 mg/L as N	Retention pond	22
Total Lead	1.20 µg/L as Pb 2.20 µg/L as Pb	Wetland basin Biofilter	5 50
Dissolved Lead	1.00 µg/L as Pb	Media Filter Biofilter Wetland basin	17 38 2
Total Zinc	19.00 µg/L as Zn	Retention pond	34
Dissolved Zinc	10.00 µg/L as Zn 19.20 µg/L as Zn	Retention pond Biofilter	9 41
Total Copper	3.0 µg/L as Cu 5.0 µg/L as Cu	Wetland basin Media filter Retention pond	4 27 27
Dissolved Copper	4.37 µg/L as Cu 5.90 µg/L as Cu	Retention pond Biofilter	9 41

1. The recommendations are based on a rationale by Dr. E Fassman; the Geosyntec and Wright Water Engineers (2008) data summary does not provide recommendations for C*.

6.4.6. Summation of Loads

The summation of loads method allows performance to be measured by calculating the ratio of all outlet loads to inlet loads.

$$\text{SOL} = 1 - \frac{Sum\ of\ inlet\ loads}{sum\ of\ outlet\ loads}$$

Or

$$\text{SOL} = 1 - \sum_{i=1}^{n} \frac{C_{inlet}V_{inlet}}{C_{outlet}V_{outlet}}$$

Where

i	duration of sample period
n	number of aliquots
C_{inlet}, C_{outlet}	inlet and outlet concentrations respectively
V_{inlet}, V_{outlet}	volumetric flow rate of inlet and outlet respectively

6.4.7. Efficiency Ratio

The efficiency ratio (ER) is defined in terms of the average Event Mean Concentration of pollutants calculated over the duration of the analysed storm.

$$\text{ER} = 1 - \frac{average\ inlet\ EMC - average\ outlet\ EMC}{average\ inlet\ EMC}$$

$$EMC = \frac{\sum_{i=1}^{n} V_i C_i}{\sum_{i=1}^{n} V_i}$$

Where

V_i	Volume of flow during period i
C_i	Concentration associated with period i
n	Total number of aliquots collected during event

6.4.8. Flow Based Variability (FBV) Curve

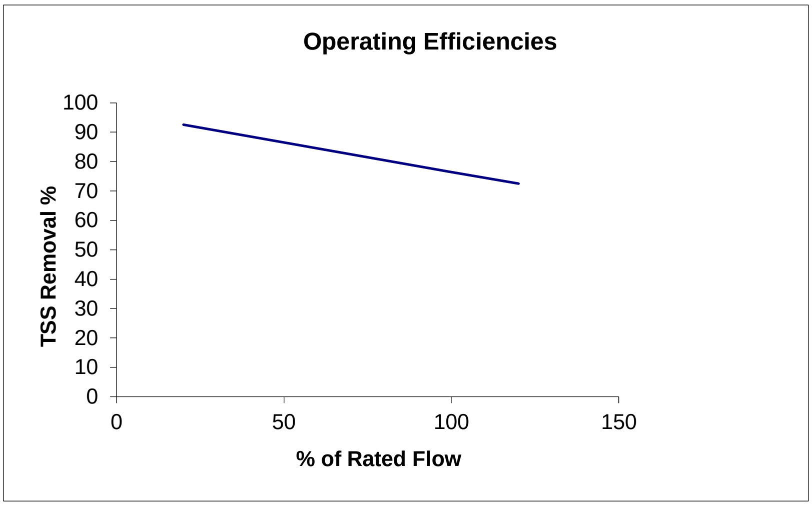
The testing protocols call for each device to establish performance efficiencies for a range of storms.

Over a sufficiently long enough period of time the ability for a device to remove pollution will lead to average effects which could be reasonably expected across different installations, however design processes will benefit from the ability for performance to be described at different flow rates, particularly where continuous simulation design techniques are employed, or where a device is being recommended for a sensitive application where a greater resolution of performance is required.

Where data is available and lends itself to presentation in a FBV format this is desirable. If direct flow versus performance data is not available there may be a case to consider how controlled tests (e.g. laboratory) may be used to augment valid field data to generate curves which can then be included in design guidance.

In developing a FBV curve a line of best fit which describes the performance claim should be produced for the entire curve or for any discrete part. This line of best fit must have a corresponding correlation co-efficient of greater than 0.9.

Figure 6-2 – Example of FBV curve



Other forms of the curve could be used that adjust for device scalability such as the volumetric loading rate (Lps/m^3).

6.4.9. Event Mean Concentration and Mass Discharge Variability

The event mean concentration (EMC) and Mass Discharge variability are required to verify the ability of the device to manage large variability in EMCs and mass discharges.

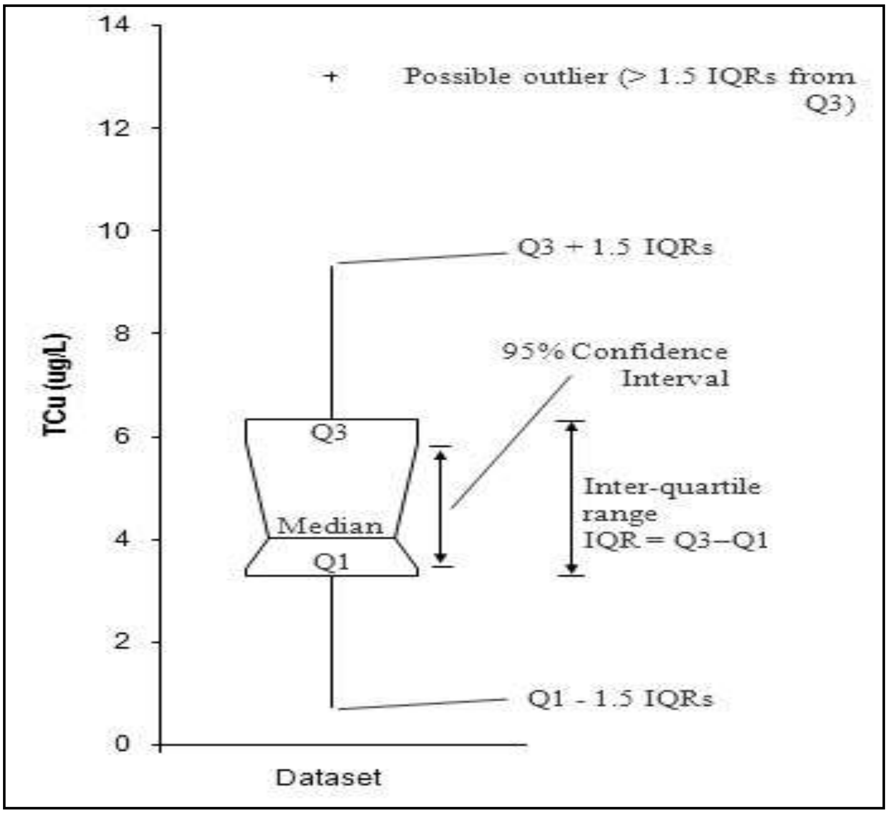
Box and whisker plots should be prepared for influent and effluent EMCs as well as mass loads. The number of EMCs and mass loads contributing to each distribution should be clearly indicated.

The following explanation of a box and whisker plot is an excerpt from Geosyntec and Wright Water Engineers (2009):

"Box plots (or box and whisker plots) provide a schematic representation of the central tendency and spread of the data. A standard box plot consists of two boxes and two lines. The lower box expresses the range of data from the 25th percentile (1st quartile or Q1) to the median of the data (50th percentile, 2nd quartile, Q2). An upper box represents the spread of the data from the median to the 75th percentile (3rd quartile or Q3). The total height of the two boxes is known as the interquartile range (Q3 – Q1). A "step" is 1.5 times the interquartile range. Two lines are drawn from the lower and upper bounds of the boxes to the minimum and maximum data points (respectively) within one step of the limits of the box. Asterisks or other point symbols are sometimes used to represent outlying data points. Some statistical packages, including stand-alone software and third-party spreadsheet extensions, also include the confidence interval about the median as notches in the boxes about the center line or can be customized to include specific data percentiles (e.g., 5th, 10th, 90th, and 95th)."

The above explanation is illustrated in **Figure 6-3.**

Figure 6-3: Example Box and Whisker Plot with Explanation of Terms from Geosyntec and Wright Water Engineers (2009)*



* While the Y-axis label indicates total copper (TCu) expressed as ug/L, the procedure is equally applicable for TSS and many other water quality parameters.

6.5. Statistical Significance Testing

Statistical significance testing of differences between inflow and outflow Event Mean Concentrations and Mass Loads is required. This significance testing determines whether the difference is too large to have occurred by chance or too small such that it is insignificant.

The selection of the appropriate statistical significance test depends on the distribution and size of the data sets. For most water quality results, the distribution is usually log-normal, with the exception of some constituents such as pH (Pitt et al, 2005). Hence, statistical testing should be performed on log-transformed data, if appropriate.

The statistical significance testing on influent and effluent data sets should be tested with the following tests, as applicable:

1. Sign Test

2. Wilcoxon-Mann-Whitney Rank-Sign Test

3. Paired t-Test

Nonparametric tests (tests 1 & 2) are only needed if the data is not normally distributed (even after log transformation). If the data is normally distributed, only the paired t-test (test 3) applies. If the data is not normally distributed even after transformation, the paired t-test is not valid. The tests should be performed on the difference between the influent and effluent data sets.

6.6. Reporting Scour

The effects of scour shall be reported if the device is an online device. Alternatively, the claimant can provide evidence that the magnitude of scour is negligible in the device The effects of scour shall be hydro-dynamically modelled; otherwise an alternative option of demonstrating scour effects should be developed as part of the QAPP and reported accordingly.

Table 7- Minimum data and qualifying event requirements for assessment

Requirements	Field Evidence Criteria	(indicative) Full Scale Laboratory Evidence Criteria[2]
Sampling Events		
Type of Event	Rainfall Events[3]	Test Runs[4] [5]
Minimum Number of Events	Statistically relevant to achieve 90% confidence interval to any required maximum number of events to achieve this. Minimum determined by defensible statistical method (e.g. ANOVA, t-test) that examines influent and effluent pairs and must be at least 8 storms and demonstrate 90% confidence interval. This may vary between target pollutants (based on catchment variability). In this event, statistical analysis can be undertaken separately for each species of interest.	15 (minimum 3 at each flow rate)- see minimum rainfall depth below for Laboratory Evidence Criteria.
Minimum Rainfall Depth	Total event rainfall depth ≥2mm	3 tests each at a constant flow rate of 25, 50, 75, 100, and 125 percent of the treatable flow rate; (for TSS) the device should be loaded with an initial sediment loading of 50% of the unit's capture capacity
Minimum/ Maximum Storm Duration/ Volume	Include a minimum sample of at least 2 storms up to 1 hour. Importantly, the minimum storm event should relate to the hydrograph and capture rising and falling limbs.	
Recommended Inter-event Time	72 hours. [6]* (see discussion points following)	
Device Size	Full Scale (where a 'family' of devices are being included as part of the claim sizing relationships must be provided for evaluation along with any basis of justification).	Full Scale (with 'family' sizing relationships being provided if subsequent field test as are being contemplated)
Runoff Characteristics	Target pollutant profile of influent and effluent	(for TSS) Particle size distribution of influent and effluent otherwise Target pollutant profile of influent and effluent

[2] These are indicative and are considered a minimum to justify field testing under local conditions. Where field tests have been previously approved the laboratory testing should be considered as supporting evidence and the veracity of claims should be considered against field test criteria providing all supporting evidence (e.g. QAPP) is provided and adhered to.

[3] Must not Include Controlled Field Tests. See glossary for the definition of controlled field tests.

[4] Includes Controlled Field Tests. See glossary for the definition of controlled field tests.

[5] from 1 site or minimum of 8 per site if more than one site is used

[6] Interevent time (ADP) will be dependent on sampling practicalities and catchment pollutant generation. Shorter ADP events may be considered where influent concentrations are above detection limits. Including minimum qualifying concentrations and aliquot collection will impose a limitation on events that can be included in analysis, but if samples are collected, their analysis and/ or omission should be disclosed for completeness of data presentation.

Runoff Volume or Peak Flow	Runoff at least 3 events should exceed 75% of the design water quality volume/ treatment flow rates of the design and at least 1 event where the device is operating at 100% or under bypass conditions.	See rainfall depth comments above
Recommended Runoff Range	Qualifying storms should include a range of events containing 0.5 to 1.1 times the peak flow and capable of being extrapolated to develop a performance curve ranging from 0.25 to 1.25 times the design flow.	As per Minimum/ Maximum Storm Duration. Device cleanout between test runs should be documented and a residual retained (and justified) to reflect any subsequent operational practicalities.
Sampling Procedures and Techniques		
Automated Sampling	Composite samples on a flow weighted basis	
Minimum Number of Aliquots	80% of field test collections should a minimum of 8 per event[7]. Notwithstanding aliquots should be collected to provide hydrograph coverage of rising and falling limbs.	
Hydrograph coverage	Indicative 50% (as a minimum the rising and falling hydrograph components should be included in testing) Dependent on catchment and rainfall patterns, multiple peaks should be accounted for (at least 1 occurrence).	Information outlining criteria for establishing laboratory testing and relationship with real world storm events

[7] Aliquot collection is determined by sample device collection rate and storm duration. For more intense, shorter duration storms a reduced number of aliquots may result. The protocol adopts a practical approach to ensure shorter duration events are able to be included in analysis and evaluation to achieve statistically robust outcomes.

Manual Sampling	Only for constituents that transform rapidly, require special preservation or adhere to bottles, or where compositing can mask the presence of some contaminants through dilution. See Section 10 for details.	
Sampling Location	As identified and agreed in the submitted QAPP.	
Analytical Methods	Various (refer to agreed EPA reference document) and/or Standard Methods (for organic, inorganic and biological analysis as required)	
Chemical and Physical analytes	As identified and agreed in the submitted QAPP.	
Minimum and maximum (influent) pollutant concentrations for qualifying events	These should be identified and agreed in the submitted QAPP. The mean concentration for 80% of samples collected for each pollution species must fall within 1 standard deviation of the expected catchment mean concentration. The mean concentration of all samples collected should fall within 2 standard deviations of the expected catchment mean concentration Influent concentrations below recommended catchment minimum concentrations (e.g. C*) should be excluded form qualifying events	
Requirements	**Field Evidence Criteria**	**Full Scale Laboratory Evidence Criteria**
Flow Measurement Location	Inlet, Outlet and Bypass*, as applicable. Based on relevant accepted measurement protocols for flow type (e.g. open channel, in pipe)	
Precipitation Measurement	Automatic rain gauge (pluviometer)	N/A
Recording Intervals	5 minutes or less	
Recording Increments	No greater than 0.25mm	
Rain Gauge Calibration	Twice during verification period	
Data Analysis and Reporting		
Performance Indicators	Based on the Performance Claim stated in Detailed Evaluation Report. (Can include but not limited to TSS, Metals, TPH, TP & TKN) The target pollutants and testing rationale must be described in the QAPP. Where a device is claiming total reductions of a particular pollutant, it is not necessary to include speciation. If speciation is not undertaken then reductions of sub-species cannot be claimed.	

Performance Indicators Calculation	1. Concentration Removal Efficiency (CRE) (See Section 6.4.3) (Arithmetic average and median. If difference is 10% or greater, inspect data set closely) 2. Mass Removal Efficiency (MRE) (See Section 6.4.4) (Arithmetic average and median. If difference is 10% or greater, inspect data set closely) 3. Relative Achievable Efficiency (RAE) (See Section 6.4.5) (Arithmetic average and median. If difference is 10% or greater, inspect data set closely 4. Summation of loads (SOL) (See Section 6.4.6) (Arithmetic Average and median. If difference is greater than 10% inspect dataset closely) 5. Efficiency Ratio (ER) (See Section 6.4.7) (Arithmetic Average and median. If difference is greater than 10% inspect dataset closely) 6. Flow Based Variability (FBV) (See Section 6.4.8), including a plot of one of the above performance measures against the 25, 50, 75, 100 and 125 percent of the treatable flow rate. Provide details on the selected curve and the associated R^2 value.	Individual removal efficiency of each test run; average runs for each operating rate; average for all runs and calculated for at least three Field Performance indicators.
Performance Variability Schematics	Box and Whisker Plot	
Statistical Significance Testing	Log-transformed inlet and outlet paired samples at 90% confidence level.	
Sizing Methodology	A sizing methodology must be provided that allows an evaluation of performance of other devices in a ‘family’ to be reviewed. This should include relationships established under defensible theoretical/ modelled conditions or testing undertaken under either field or laboratory conditions.	