Sewage Treatment

water e-journal
Online Journal of the Australian Water Association

ISSN 2206-1991
Volume 6 No 1 2021
https://doi.org/10.21139/wej.2021.005

Estimating ADWF at Sewage Treatment Plants

Determining the best technique for calculating average dry weather flow helps improve facilities' management

D de Haas, S Ng, N Dahl, D Baulch

ABSTRACT

Estimating and understanding Average Dry Weather Flow (ADWF) is fundamental to the planning, design, and operation of sewage treatment plants (STPs). This paper reviewed methods for estimation of ADWF, in four general groups: Rainfall-based; Equivalent person (EP) based; Basic statistical (Percentiles); and 'Novel'. The 'Novel' methods identified were: Histogram/ Mode; Antecedent Precipitation Index (API); Ratio of Short Term and Long-Term Moving Averages; K-means Clustering; Diurnal Profile Smoothing; and Kernel Density Estimation. EP-based methods were not considered useful because they shift the uncertainty from rainfall and/or flow data to population and/or loading data. The other methods were tested using datasets for two STPs of similar size (ADWF approximately 1.2 to 1.3 ML/d) in northern New South Wales, one of which is more prone to wet weather inflow/ infiltration (I/I). On balance of simplicity and performance against more complex methods, we recommend the Histogram/ Mode and/or the Percentile methods for routine reporting. For larger and more complex assignments (e.g., design projects, planning studies), it is recommended that one or more of the alternative high-performing methods described in this paper (e.g., Ratio of moving averages; Kernel Density Estimation) be employed for ADWF checks. Relatively large datasets (at least one year of daily flow totals) should be used and the results compared against the estimates from simpler methods.

Keywords: average dry weather flow rate, rainfall, sewage treatment plants

INTRODUCTION

Historically, both locally and internationally, engineers and managers of water utilities have applied various methods to determine the average dry weather flow (ADWF) for sewage treatment plants (STPs). To our knowledge, at least in Australia, there is no single industry standard method to define or determine ADWF. Common methods (typically) apply some form of numerical 'filter' to the totalised daily data for STP inflow, based on concurrent rainfall records. Such methods vary in detail (e.g., number of days, rainfall amount etc. applied to the numerical filter). Furthermore, such methods are constrained by the issue of representative rainfall data to apply for STP catchments (e.g., accuracy of rainfall records, nearest geographic factors, spatial distribution of rainfall) and the variable degree to which rainfall affects STP inflows (e.g., prolonged effects of rainfall on infiltration/ inflow (I/I) in some but not all catchment sewer systems).

A review of industry 'best practice' methods for defining and calculating ADWF at STP was collated from various sources, including the following: existing environmental licenses in Australia (QLD and NSW); international governing bodies (UK Government Environment Agency and Winnipeg Water and Waste Department); and industry practice (e.g., previous approaches used by consultants, suggested alternatives by local councils, or industry associations etc.).

The methods reviewed fell into four groups:

1. Rainfall-based
2. EP based
3. Basic statistical
4. Novel

Rainfall-based methods attempt to determine which days were dry by examining historical daily rainfall records and, in many cases, also looking at rainfall on a given day along with preceding days.

EP-based methods attempt to determine dry weather flow empirically by estimating the number of equivalent persons (EP) within the catchment area and multiplying by an average wastewater production per EP.

Basic statistical methods were those that apply a very simple statistical analysis of flow data. The only example of this found in the literature review was Percentile-based estimations; however, using a histogram or mode calculation would be very similar. Percentile-based methods look at the entire set of flow data, including wet days, and estimate the average dry weather flow by taking a flow percentile (typically between the 20^{th} and 50^{th} percentile). The histogram/ mode method looks at the frequency of different flowrates and takes the most commonly occurring flowrate as an estimate for ADWF.

'Novel' methods were those that did not fall into one of the other three groups described above. In the literature review, only one method was considered novel, namely a 'ground-up' approach for very small flow systems. The average flow for household water-using devices (taps, showers, appliances etc.) is estimated and ADWF is then stochastically estimated based on the frequency and duration of use.

The aim of this paper was to compare the results of using the current industry 'best practice' methods for estimating ADWF against five novel estimation methods that seek to improve estimate performance. To our knowledge, the novel methods we selected for testing have never been applied for this purpose to STP flows on a routine basis.

METHODOLOGY

Data

Flow data from 2011 to 2020 was supplied by Byron Shire Council (BSC) for two sewage treatment plants (STP) in northern New South Wales: Ocean Shores STP (OSSTP) and Brunswick Valley STP (BVSTP). The two plants are located within a straight-line distance of 1.7 km from each other.

Additional short time interval flow rate data from 2017 to 2019 was provided for BVSTP for use in the Diurnal Profile Smoothing method, as described below.

Rainfall data was retrieved from the nearby weather stations from Bureau of Meteorology (BOM) data: Mullumbimby and Brunswick Heads Bowling Club (BOM station no. 058040 and 058103, respectively).

Overview

Investigations began with a preliminary set of estimates using current industry methods. The initial methods tested were:

- 20^{th} percentile
- 30^{th} percentile
- 50^{th} percentile
- QLD EPA SEQ Rainfall-based method
- One-week Rainfall-based method
- Three-week Rainfall-based method.

These informed the initial rating system for discerning good vs. poor estimates. To do this, the above-mentioned methods were separated into three levels of strictness: Least Strict, Moderately Strict, and Strictest. The Strictest estimates were expected to be the highest performing as they eliminated the most data for days influenced by rainfall events.

Six additional methods were proposed - one basic statistical, and five 'novel' methods - as follows:

- Histogram / Mode (basic statistical method)
- Antecedent Precipitation Index (API)
- Ratio of Short Term and Long-Term Moving Averages
- K-means Clustering
- Diurnal Profile Smoothing
- Kernel Density Estimation

Each method was evaluated and compared against the results of the Strictest estimates from the above-mentioned rating of the initial methods tested. Based on the highest performing among the initial and proposed additional methods, a 'true value ADWF' was adopted for the datasets examined from each of the two STPs.

To assess the ability of individual methods to estimate ADWF, the results for each were compared against the 'true value ADWF' for the two STPs and a relative score given. All methods were then compared in a Multi-Criteria Assessment (MCA), which scored the methods semi-quantitatively for Estimate Performance, Data Requirements, Mathematical Complexity, Parameter Complexity, and Robustness.

DESCRIPTION OF NOVEL ESTIMATION METHODS

Antecedent Precipitation Index (API) Method

The first novel method is similar to rainfall-based methods. It applies an established modelling term, namely Antecedent Precipitation Index (API). API is a running day-by-day index of moisture stored within a drainage basin (Ali, et al., 2010). The difference between API versus a simple cumulative rainfall is that API considers the nature of catchments where drying out progressively occurs during periods without rain, making recent rainfall events more impactful than earlier events. Mathematically, API takes the form of:

$$API = \sum_{t=0}^{-i} P_t k^{-t}$$

Where:

i is the number of antecedent (preceding) days considered

k is the decay constant (d^{-1})

P_t is the rainfall during a given day at time, t

t is time (days).

For this study, the values we chose for *k* and *i* were 0.9 (Ali *et al*., 2010; Kohler & Linsley, 1951) and 27 days, respectively. This approach considers rainfall over the 27 antecedent days up to and including a given present day (28 days total). The index progressively places less weighting on rainfall that occurred in earlier antecedent days, culminating in a weighting of 5% for rainfall measured on the 27th antecedent day.

We defined a dry day as any day with $API \leq 10\ mm$. Using this definition, ADWF was then calculated as the median of dry day flows.

Ratio of Short Term and Long-Term Moving Averages

The second novel method is based on the ratio of short-term to long-term moving averages of daily flows. Based on similar applications for monitoring variance within natural systems, including human fitness (Murray *et al*., 2016), this method compares the short and long-term averages to determine if flow is changing significantly or relatively stable. A ratio between the short and long-term averages close to unity (1.0) is taken as an indicator of stable flow and, by implication, dry weather flow.

This method has the benefit that it technically does not classify "dry weather" based on rainfall but rather attempts to discern baseline flows. This has the advantage that it can be equally applied in regions with different climates, including those where rainfall occurs relatively frequently, causing I/I to produce on-going contributions to average flow. Similarly, it can be applied in situations where local rainfall records either do not exist or are unreliable. By contrast, rainfall-based methods depend on reliable rainfall data and a single definition of 'dry weather' is difficult for different situations.

Arithmetic Moving Average

The simplest form of calculating moving averages is arithmetically. In this case the moving average is the sum of flowrates divided by the number of days considered. Mathematically, the formula is:

$$\overline{F} = \frac{\sum_{t=1}^{i} F_t}{i}$$

Where:

$\overline{F}$ is the moving average

i is number of days considered

t is the time (days)

F_t is the flowrate on day t

For this study, we chose *i* to be 7 days (1 week) for the short-term average, and 28 days (4 weeks) for the long-term average. The ratio of short-term average to long-term average (ϕ) is simply expressed as:

$$\phi = \frac{\overline{F}_{ST}}{\overline{F}_{LT}}$$

Where $\overline{F}_{ST}$ and $\overline{F}_{LT}$ are the short term and long-term moving averages, respectively.

To classify dry weather days, by trial and error we selected an upper bound for the ratio (ϕ) of 1.025 and a lower bound 0.976 (the inverse of the upper bound).

Exponentially Weighted Moving Average

The exponentially weighted moving average (EWMA) is a modification to the arithmetic moving average where a diminished weighting is applied to older flowrates, like the API method. The formula for the EWMA is:

$$\bar{F} = \frac{\sum_{t=0}^{i-1} F_t k^{-t}}{\sum_{t=0}^{i-1} k^{-t}}$$

Where:

$\bar{F}$ is the moving average

i is the number of days considered

t is time (days)

F_t is the flowrate on day 't'

k is the decay constant.

As before, we chose value of *i* = 7 days for the short-term average, and *i* = 28 days for the long-term average. We chose *k* to be 0.9, consistent with the API method (see above). As before, to classify dry days we chose an upper bound for the ratio of 1.025 and a lower bound of 0.976.

Ratio of Moving Averages with Step-change Limit

A modification of the Ratio of Moving Averages method was developed to limit the impact on results from large rainfall events. This modification excludes readings where the ratio of averages changes rapidly. For example, after a heavy rainfall event the short-term moving average flow spikes to a high value before returning to a lower flow. As the short-term average recedes after the high flow event, the long-term average flow rises, and the two averages will intersect at a flowrate that is higher than the ADWF. The ratio step-change limit prevents such intersections from being counted as dry days. For the step-change limit, we applied the logic that a given day (at time t) is considered dry if:

$$\Delta\phi_t < Step\ Change\ Limit$$

$$For\ t\ = 0\ to\ i-1, \qquad \Delta\phi_t = \phi_{t+1} - \phi_t$$

Where:

ϕ is the ratio of moving averages (arithmetic or exponentially weighted, see above) and t is time (days).

For this study, by trial-and-error we selected a step-change limit of 0.025.

K-means Clustering

The third novel method is a clustering approach intended to use an advanced method of classifying flowrates with the aim of isolating dry weather flows from other groupings of daily total flows. Taken from machine learning techniques, K-means clustering seeks to sort a series of observations into a number (k) of groups called clusters, thereby revealing underlying patterns. The number of clusters is specified by the user and an algorithm seeks to select centroids such that the distance from that centroid to points within its cluster is minimised. In the case of ADWF, clusters are selected such that similar daily flowrates are grouped together. A real-life analogue might be a four-cluster grouping such as: dry weather, light rain, heavy rain, and extreme weather. Figure 1 demonstrates the idea of K-means clustering.

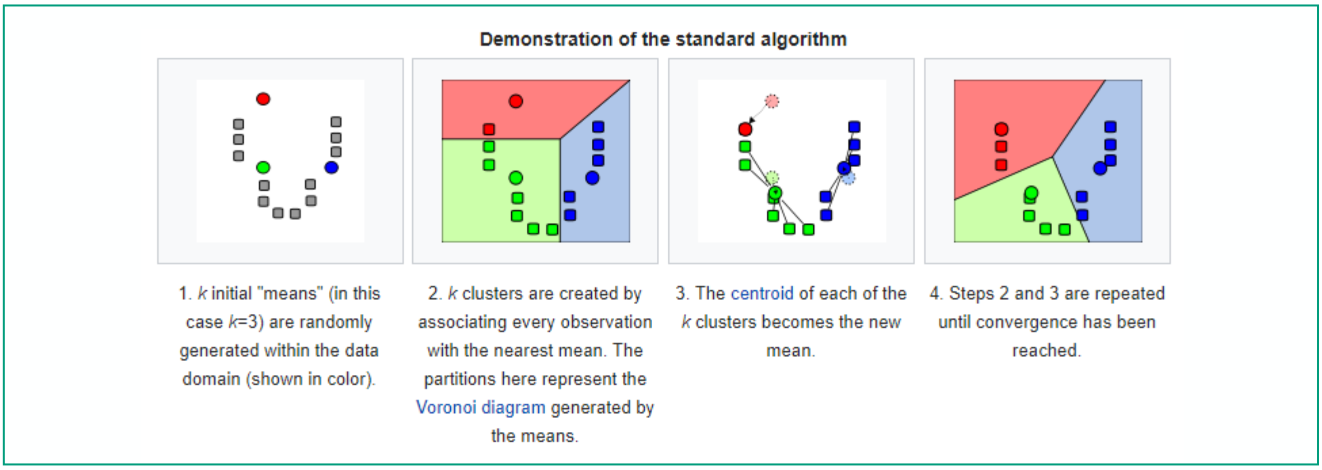


Figure 1: Diagram of K-means Clustering (Wikipedia, 2020c)

Diurnal Profile Smoothing

The fourth novel method is an advanced classification method attempting to discern base flowrates from short time-interval data (intervals of minutes or hours). The Diurnal Profile Smoothing method is predicated on identifying higher flow rates indicative of wet weather and separating these from the dry weather flow pattern of an STP. This is possible using large datasets of short time-interval flow rate to overlay plots at-weekly time spans. Removing outliers (identified wet weather data) and averaging or smoothing the dry weather weekly flow pattern, produces an estimate of the underlying base flow.

Once the underlying base flow pattern is known, ADWF can be calculated by integration. Figure 2 shows an example of the curves produced by a Gaussian Kernel Regression (the method of smoothing we have chosen).

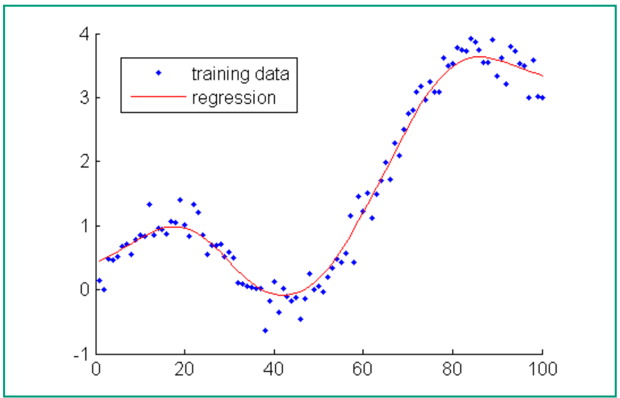


Figure 2: Gaussian Kernel Smoothing (Wikipedia, 2020b)

Kernel Density Estimation

The fifth novel method is a modified histogram/ mode approach that attempts to create a continuous distribution from flow data rather than the discrete formulation of a histogram. The translation to average dry weather flow is the same as for a regular histogram, namely that the most common flowrate is likely to be a good estimate of ADWF.

Kernel density estimation produces a continuous distribution by plotting all the data points (flowrate) onto the X-axis and assigning a distribution function for each point, called a kernel. In the case of Figure 3, a Gaussian distribution has been assigned and centred at each data point (dashed red lines). The final distribution is obtained by summing the values of the individual kernel at continuous x values, resulting in peaks for ranges with many data points and troughs for ranges with few.

Kernel density estimation has two advantages. Firstly, it produces a clearer picture when there are relatively few data points, compared with a histogram. Secondly, it provides weighting to adjacent data points so that a more representative estimate between data points and at extremes is created, which can improve performance with smaller and/or skewed data sets.

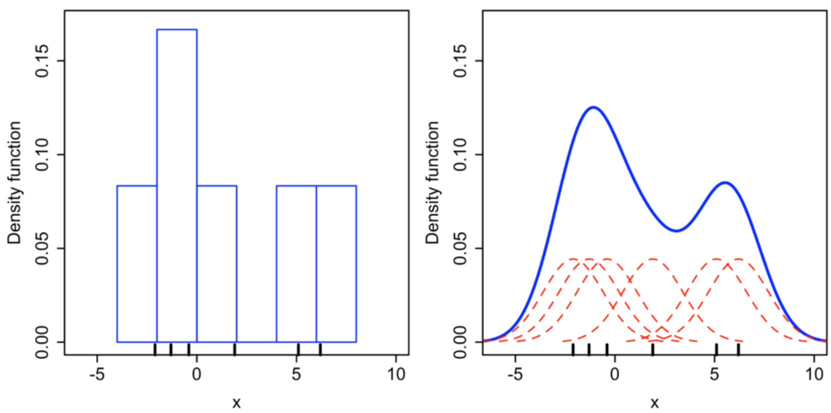


Figure 3: Histogram vs Kernel Density Estimation (Wikipedia, 2020a)

RESULTS

Basic Statistical Methods

A summary of the results for the Percentile-based methods is given in Table 1.

Table 1 Summary of Percentile-based Estimates

Percentile	OSSTP (ML/d)	BVSTP (ML/d)
20th	1.25	1.19
30th	1.29	1.26
40th	1.33	1.36
50th	1.39	1.46

It is useful to compare the percentile results with the mode and traditional histogram for the same datasets. Using the original flow datasets (without editing i.e., including potential outliers) and taking the mode as an estimate for the ADWF, results in a value of 1.29 ML/d and 1.16 ML/d for OSSTP and BVSTP respectively. As an example, the histogram for OSSTP is shown in Figure , where the probability plot and 20th percentile are also plotted, demonstrating similarity between the methods.

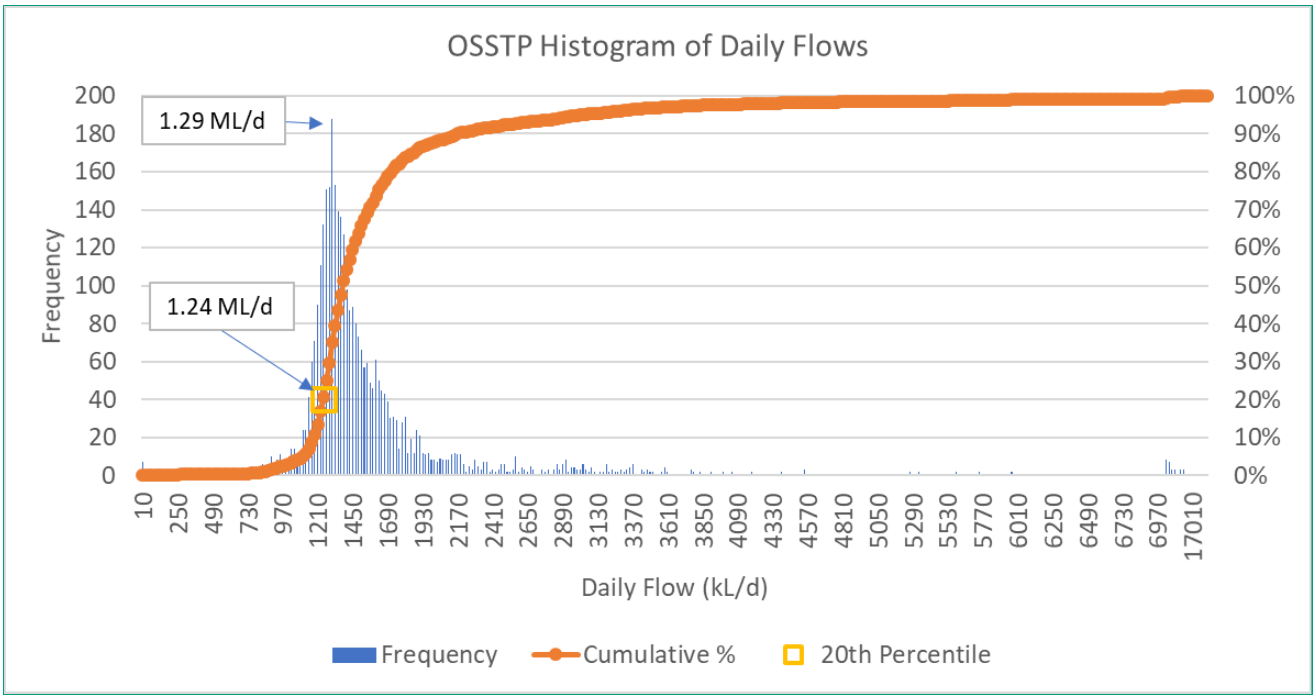


Figure 4: OSSTP Daily Flow Histogram

RAINFALL-BASED

The calculated ADWF was analysed for sensitivity to the number of preceding days considered, and the average rainfall over that period (i.e., a given day and its preceding days). Refer to Table 12 in *Supplementary Information* for the detailed results. It was concluded that considering at least six preceding days and allowing up to 7 mm cumulative rainfall over that day and its six preceding days, gave a sufficiently strict definition of a dry day. This amounted to an *average* rainfall of up to 1 mm/day *over seven consecutive days*. It enabled an estimate of ADWF to within a margin of 10% of the results produced by the strictest parameters tested. The strictest parameters tested were 27 preceding days and 0 mm/d of *average* rainfall (i.e., 0 mm rainfall on any given day *and in aggregate over 28 days*). By way of illustration, the ADWF estimated by allowing 1 mm/d average rainfall over one week (i.e., either <7 mm in aggregate over seven consecutive days), or three weeks (i.e., <21 mm in aggregate over 21 days) are both shown in Figure 5 and Figure 6 below for OSSTP and BVSTP, respectively.

The Rainfall methods can be compared with the Percentile method. The ADWF estimated from the rainfall using the stricter definition of a dry day (<21 mm in aggregate over 21 days) lay between the 20th and 30th percentiles, whereas that estimated using (<7 mm in aggregate over 7 days) lay between the 30th and 50th percentiles (see Figure 5 and Figure 6). The ADWF estimated from rainfall using a method applied in a recent Queensland STP environmental license (QLD EPA, 2020) was higher than the 50th percentile (i.e., significantly higher than that from the stricter rainfall definitions applied here - see above).

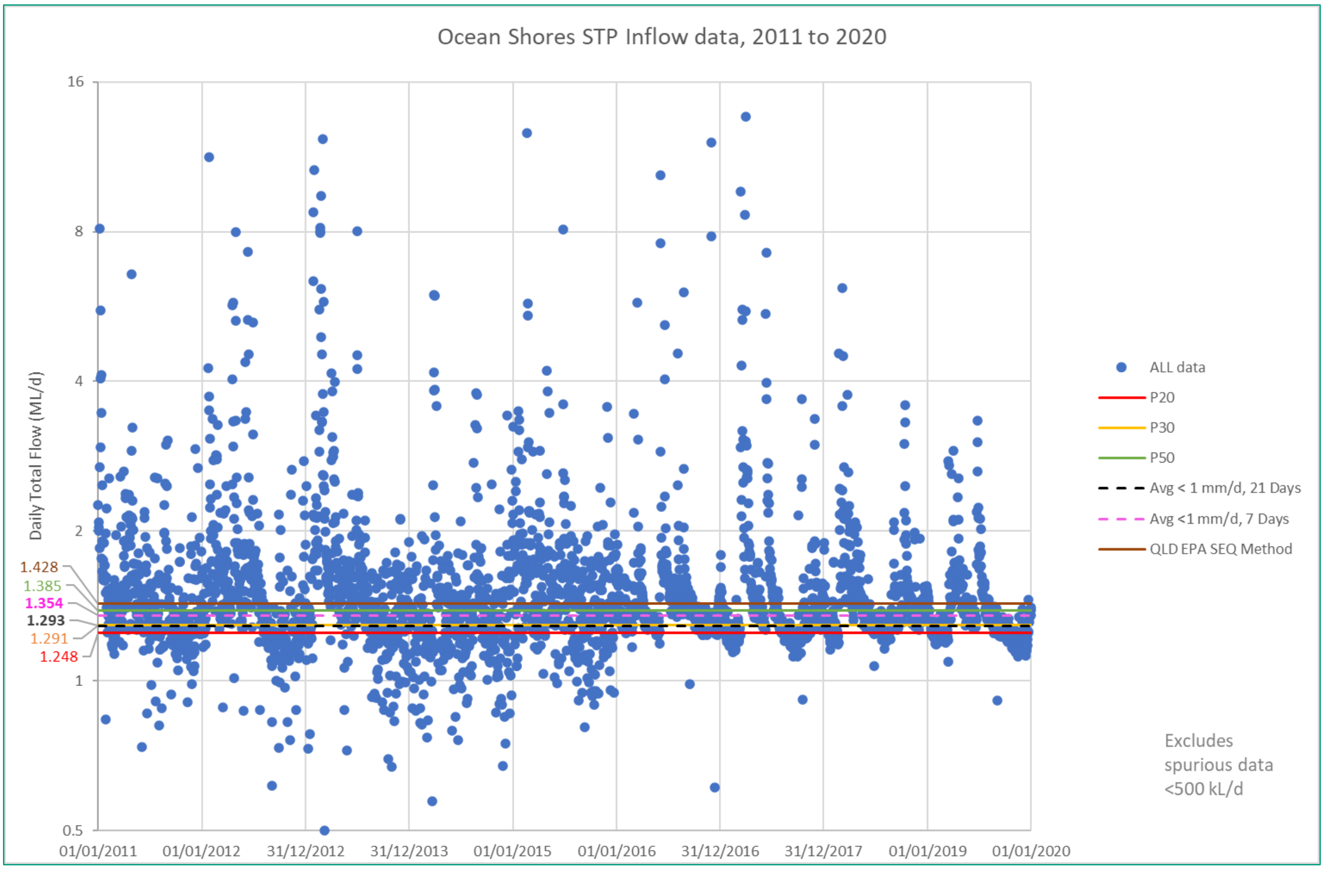


Figure 5: Ocean Shores STP Comparison of Existing Methods

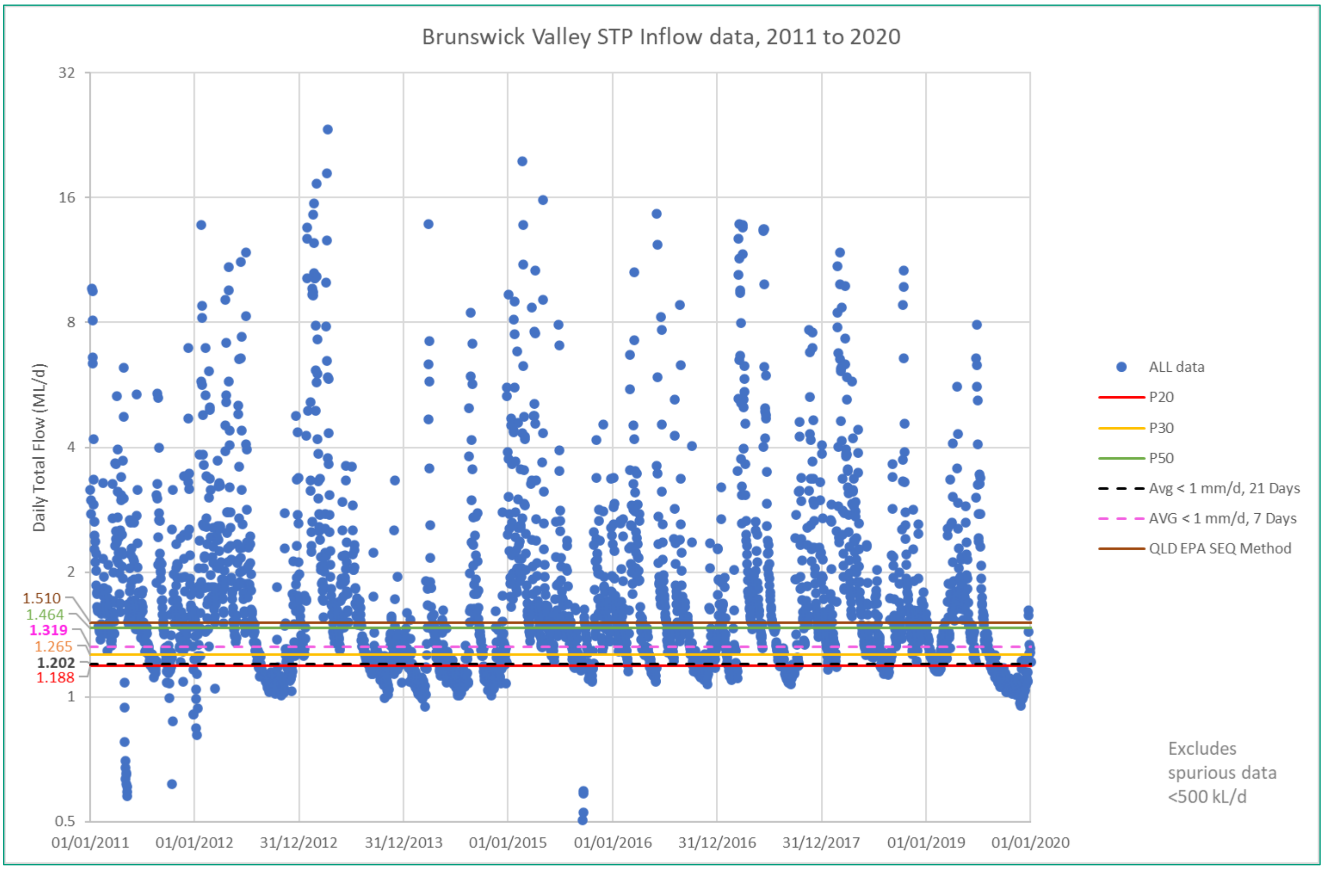


Figure 6: Brunswick Valley STP Comparison of Existing Methods

NOVEL ESTIMATION METHODS

Antecedent Precipitation Index (API) Method

The results for the API method are summarised in Table 2. Like the rainfall method results (see above), the ADWF estimates lay between the 20th and 30th percentiles (Table 1).

Table 2: ADWF estimates by API method

Site	Estimated ADWF from median on dry days (ML/d)
OSSTP	1.29
BVSTP	1.19

Ratio of Short Term and Long-Term Moving Averages

Taking the ADWF as either the mean or median of flows classified as dry by the ratio of moving averages method yields the results in Table 3.

Table 3: ADWF Estimates by Ratio of Moving Averages method

Method of moving averages	ADWF on dry days (ML/d)			
	OSSTP		BVSTP	
Method of calculating ADWF on dry days	Mean	Median	Mean	Median
Arithmetic	1.36	1.30	1.35	1.19
Arithmetic + Step-change Limit	1.32	1.29	1.21	1.16
EWMA	1.39	1.31	1.39	1.18
EWMA + Step-change Limit	1.35	1.30	1.23	1.16

Looking at the ADWF calculated from the mean flows on dry days, the arithmetic moving average and EWMA methods vary by less than 0.05 ML/d without the step-change limit, and by less than 0.03 ML/d with the step-change limit. With the parameters chosen, the two methods for calculating moving averages are almost indistinguishable in terms of ADWF estimated.

However, the step-change limit had a significant effect on the ADWF estimate (from the mean flows on dry days), particularly for BVSTP where it lowered the estimate from 1.39 to 1.23 ML/d (i.e., by a margin of 11%). Figure 7 illustrates the residual effects of wet weather impacts on flow rates for dry days, as defined by the moving average methods for the two STPs. Anecdotally, BVSTP is more prone to high wet weather flows than OSSTP.

Noting the impact of the step-change limit on the estimated ADWF using the moving average methods, as an alternative, we found that the effect of peak weather events could be reduced in the calculation by replacing mean flow on dry days with the median flow (on dry days). A comparison can be made from the results in Table 3. The median consistently predicts a lower ADWF than the mean and gives very little change in the result when altering or completely removing the step-change limit. With or without the step-change limit, the ADWF estimate changed by 0.02 ML/d or less (a margin of <2%).

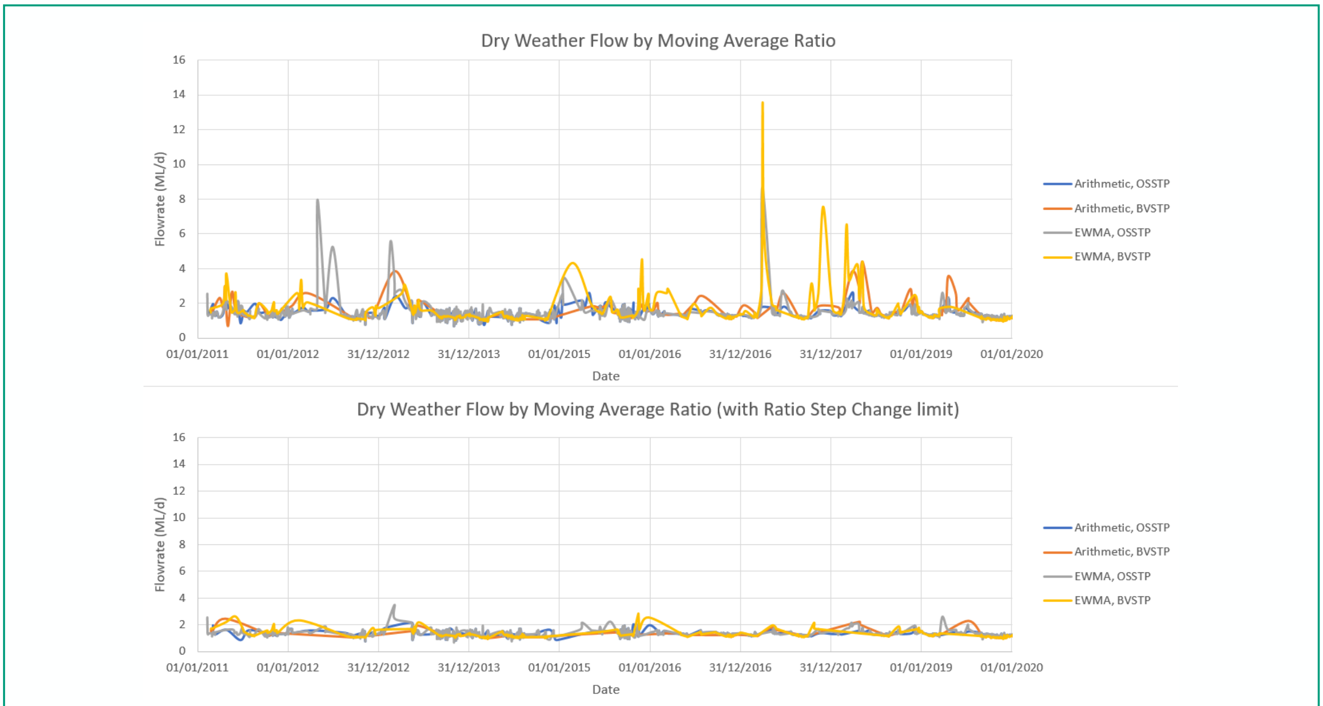


Figure 7: Dry Weather Flow as defined by Ratio of Moving Averages (BVSTP and OSSTP)

K-means Clustering

Data was tested from two to ten clusters (k=2 to k = 10) and the number of clusters that best grouped the data was used (as measured by the Silhouette coefficient). The Silhouette Coefficient (minimum is -1, maximum is +1) is dimensionless. It is used as a measure of clustering for a dataset with a larger value (closer to +1) indicating better clustering. The Silhouette Coefficient for the two datasets in Figure 8 ranged from 0.6 to 1.

The results of K-means clustering on OSSTP and BVSTP are displayed below in Figure 8 for k = 2, as this gave the highest Silhouette Coefficient. The final Silhouette Coefficients were 0.97 and 0.95 for OSSTP and BVSTP, respectively.

ADWF was estimated by the centroid of the lowest cluster (red square in Figure 8, and the results are shown below in Table 4.

Compared with results for other methods (see above), the K-means Cluster estimates of ADWF (Table 4) were relatively high and likely to be over-predicting ADWF. This is due to a relatively uniform distribution of flowrates in the datasets (see Figue 8). The underlying cause is the I/I pattern, influenced largely by climate, location and the nature of sewer catchments served by the STPs. It was noted that this method performed well when tested on a STP in North Queensland where the flow pattern is more multimodal. Multimodal flow patterns would be characterised by two or more *distinctly different* sets (or clusters) of flowrates, for example, a distinct wet season vs. dry season, which tends to be characteristic of North Queensland.

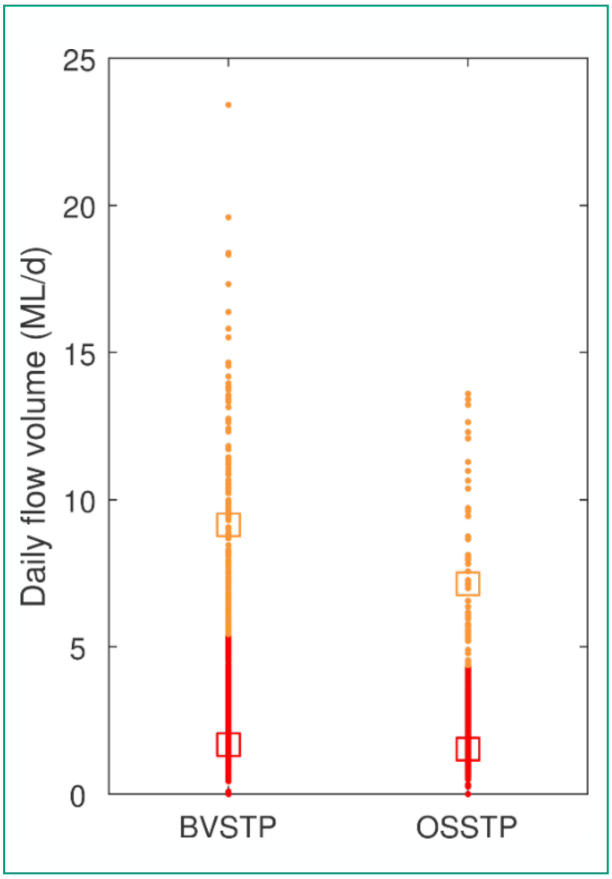


Figure 8: K-means Clustering results for OSSTP and BVSTP

Table 4: K-means Clustering Results

Site	Estimated ADWF
OSSTP	1.53
BVSTP	1.68

Diurnal Profile Smoothing

Plots of 3-hour and 0.5-hour aggregations of diurnal flow rates using the Diurnal Smoothing Method are shown below in Figure 9 and Figure 10 for BVSTP data. This method was not tested for OSSTP data, due to study constraints. The ADWF results for both calculations are presented in Table 5.

Table 5: Diurnal Profile Smoothing Results

Site	Estimated ADWF (3 Hour data aggregation) (ML/d)	Estimated ADWF (0.5 Hour data aggregation) (ML/d)
BVSTP (with outliers, for reference only)	2.27	2.27
BVSTP (without outliers)	1.63	1.15

Table 5 shows that the results are sensitive to the time interval chosen. To obtain an accurate estimate of ADWF, namely that aligned with existing methods, the time interval had to be decreased to 0.5 hours. The ADWF estimate was 1.15 ML/d, which aligns closely with that from other methods (e.g., the Percentile or Rainfall-based methods – see Figure 6). The inclusion of outliers at the same time interval gave a result of 2.27 ML/d (Table 5), demonstrating the importance of removing outlier weekly profiles when using this method.

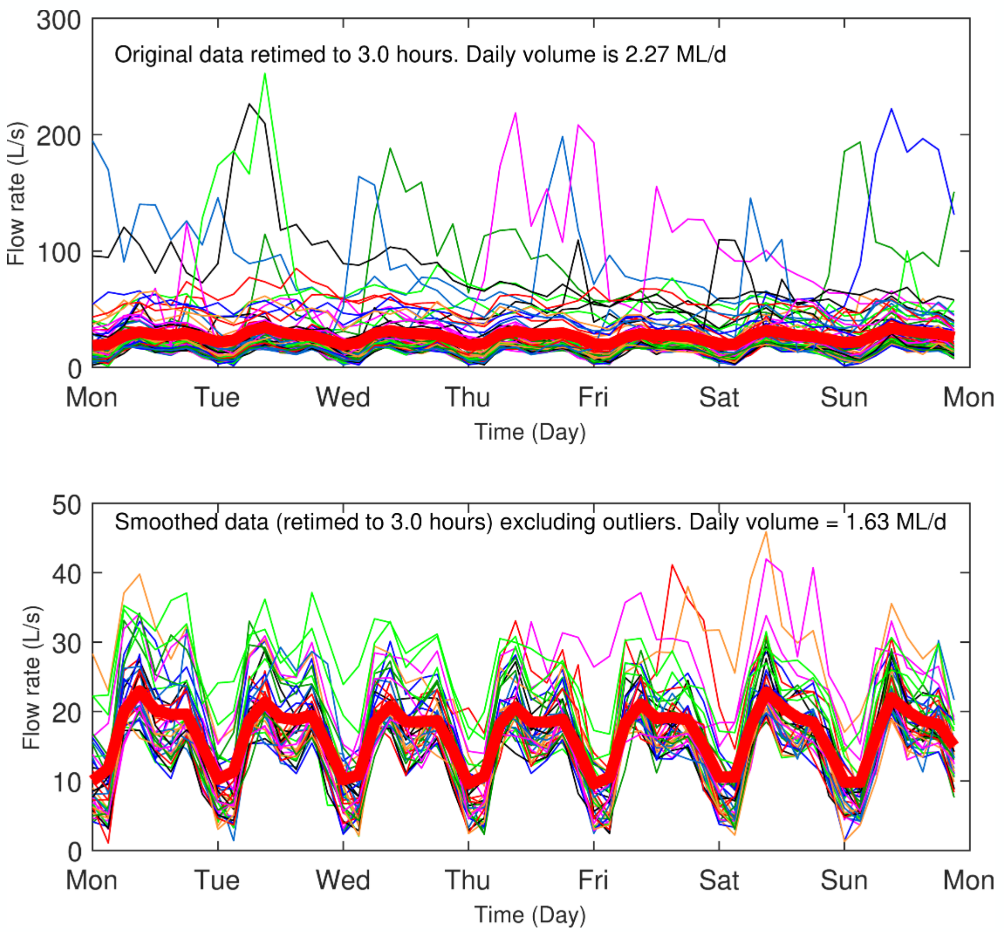


Figure 9: BVSTP Diurnal Smoothing (3 Hour data aggregation)

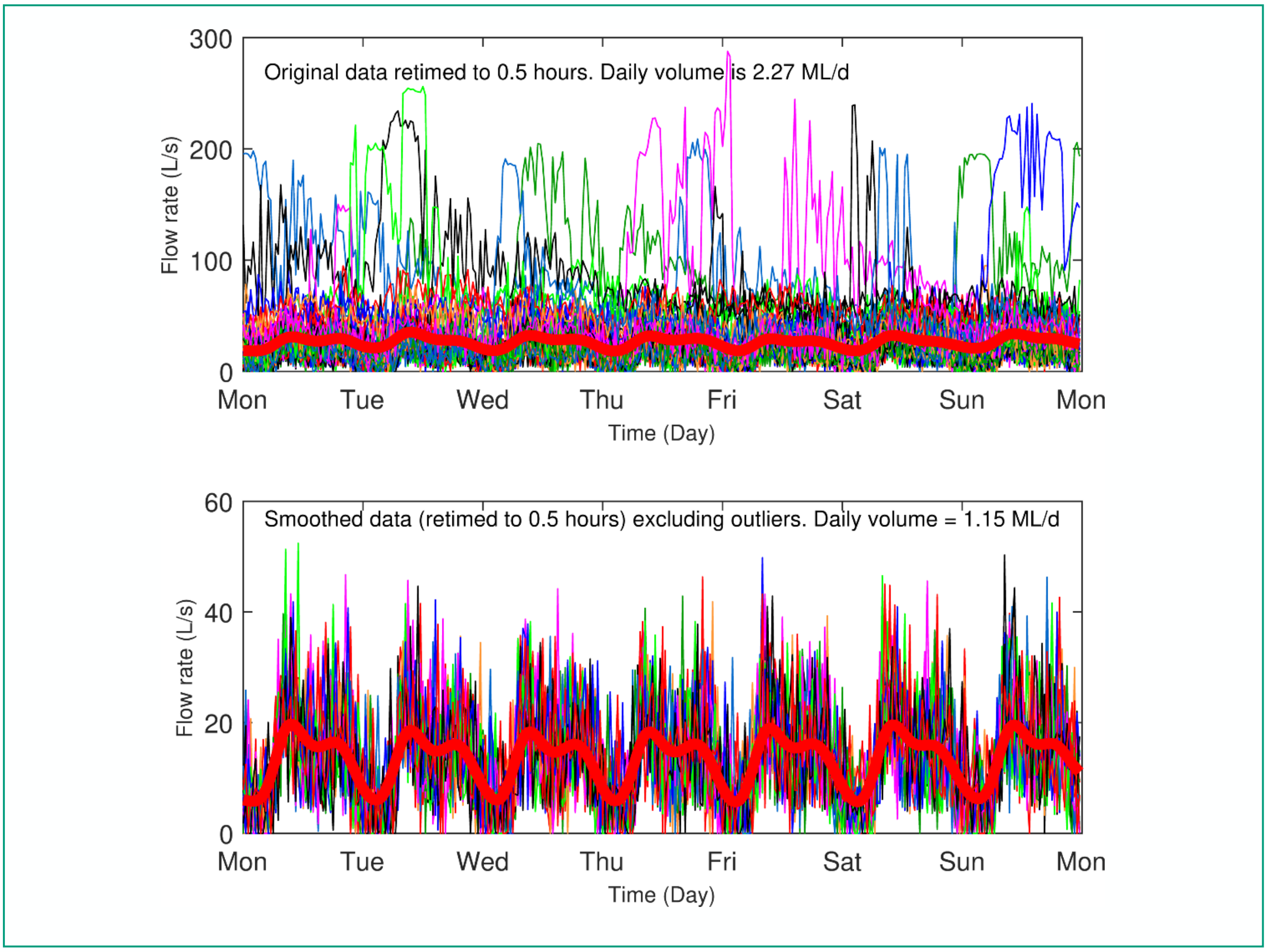


Figure 10: BVSTP Diurnal Smoothing (0.5 Hour data aggregation)

Kernel Density Estimation

Kernel density estimation was performed on both OSSTP and BVSTP using all daily flow data, including potentially spurious data. The distributions are shown below in Figure 11. It can be seen in the figure that the histogram is jagged while the Kernel density distribution is smooth. For BVSTP, the mode is at 1.29 ML/d while the Kernel density estimate is 1.22 ML/d, highlighting the advantage of this method.

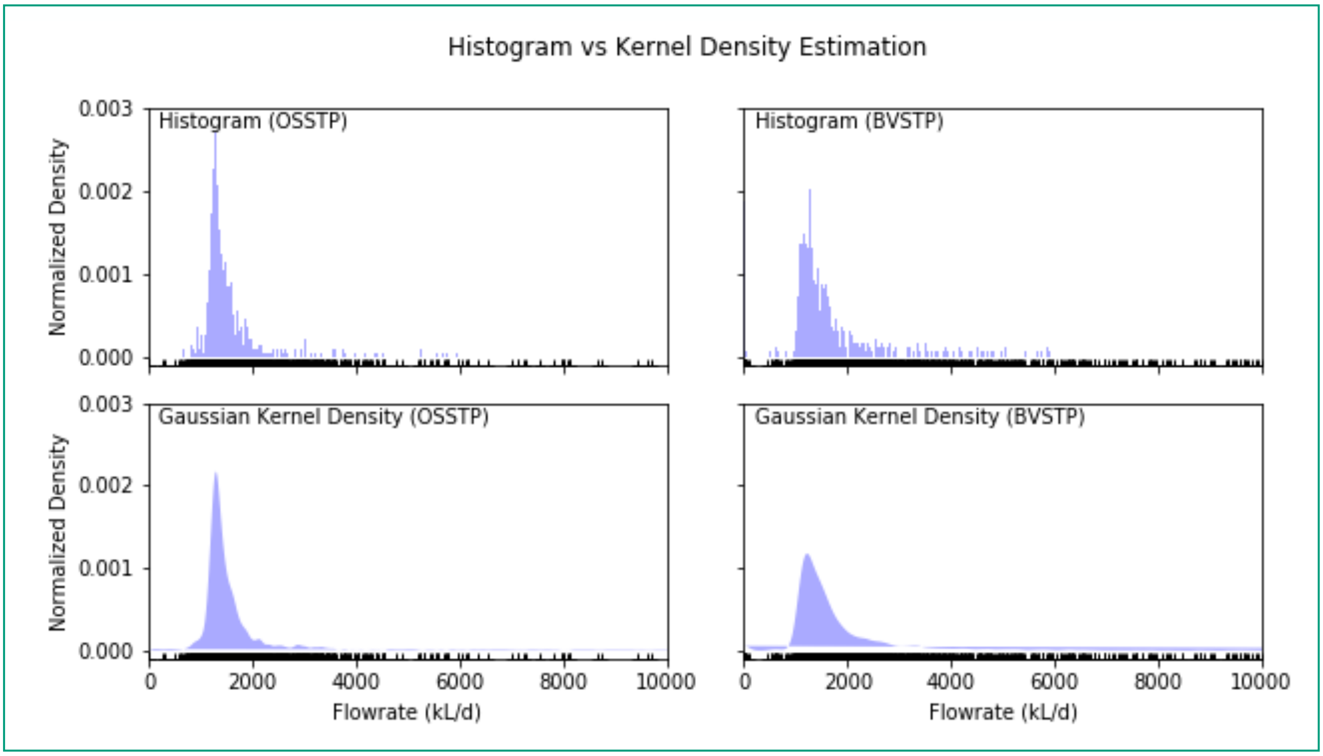


Figure 11: Kernel Density Estimation Results

Using the peak of the continuous curves (Gaussian Kernel Density distributions) as the estimate for ADWF for each plant produces the results in Table 6. The results for both STPs align well with the previous high performing methods (e.g., ratio of moving averages), as well as the existing 'Strictest' methods (e.g., Percentile or Rainfall-based).

Table 6: Kernel Density Estimation Results

Site	Estimated ADWF (ML/d)
OSSTP	1.30
BVSTP	1.22

DISCUSSION

A multi-criteria assessment (MCA) approach was used to review and compare all the ADWF estimation methods (existing and novel) developed and/or tested in this study.

Estimation Classifications

Rating estimates of ADWF as 'good' or 'bad' is to some extent subjective, given that the true ADWF value is uncertain and the estimation methods differ in complexity, which is a user judgement. To meaningfully discuss the results of different estimates, we applied a classification system based on 'strictness'. Based on preliminary results from existing methods commonly used in industry and the literature (i.e., Flow percentile and Rainfall-based methods), we adopted three nominal classifications that hinge on the definition of 'dry days': 'Least Strict', 'Moderately Strict', and 'Strictest', as summarised in Table 7 below.

Table 7: Classification of existing methods used for assessment in this study

Classification for ADWF definition of 'dry days'	Methods	Comments
Least Strict	Existing QLD EPA SEQ Rainfall-based method 50th Percentile	For both OSSTP and BVSTP, the QLD EPA SEQ method and the 50th Percentile produced similar ADWF estimates (approx. 1.39 ML/d for OSSTP; 1.46 ML/d for BVSTP). However, both overestimated ADWF compared with the 'stricter' versions of the same methods.
Moderately Strict	One-week Rainfall-based method (<1 mm/d average rainfall over seven days) 30th Percentile	Both methods gave ADWF estimates that were somewhat lower than the Least Strict methods (1.29 to 1.37 ML/d for OSSTP; 1.27 to 1.34 ML/d for BVSTP).
Strictest	Three-week Rainfall-based method (<1 mm/d average rainfall over 21 days) 20th Percentile	The 20th Percentile method gave the lowest estimate for both STPs, followed by the Rainfall-based method (<1 mm/d average rainfall over 21 days). These two methods produced similar ADWF estimates (1.25 to 1.30 ML/d for OSSTP; 1.19 to 1.20 ML/d for BVSTP).

Multi-Criteria Assessment (MCA)

To assess the ability for each method to accurately estimate ADWF, a 'true value ADWF' needed to be established for each STP. The 'Strictest' method estimates were approximately 1.30 ML/d for OSSTP and 1.20 ML/d for BVSTP. Whilst most of the novel methods gave results that agreed closely with this estimate for OSSTP, a number produced lower estimates for BVSTP.

The adopted 'true' ADWF value was chosen based on the average of the high performing methods (i.e., methods giving estimates closest to those using the 'Strictest' methods classified in Table 7 and in closest agreement with each other), namely: Histogram/ Mode, API, Ratio of Moving Averages, Diurnal Smoothing, and Kernel Density Estimation.

The adopted 'true' ADWF values were 1.30 ML/d for OSSTP and 1.18 ML/d for BVSTP.

Methods were then scored (Table 8) based on their ADWF estimates for both STPs, using the sum of the difference between the method estimates and the adopted 'True ADWF' values.

Since estimation performance is similar among many of the high-performing methods, additional criteria were considered in the MCA. These criteria were data requirements, mathematical complexity, parameter complexity, and robustness (e.g., sensitivity to choice of parameters; sensitivity to data errors). The scoring criteria for these are also listed below in Table 8. For all criteria, a low score value is better.

The scores against all criteria for characteristics were summed and weightings applied to determine the overall score for a given method. The weightings are given in Table 9.

The results of the MCA are presented in Table 10.

Table 8: MCA Scores for Criteria

Score	**Performance for ADWF estimation**	**Absolute (±) Difference between ADWF Estimate and 'True ADWF' (ML/d)**
1	Excellent	<=0.1
2	Good	>0.1 to 0.2
3	Moderate	>0.2 to 0.4
4	Poor	>0.4
Score	**Data requirements**	**Descriptors**
1	Very Low	Daily flow data (small sample size)
2	Low	Daily flow data (large sample size)
3	Moderate	Daily flow data + rainfall data
4	High	Hourly and/or minute time-step flow data
Score (Note 1)	Mathematical complexity	Descriptors
1.3	Low	Basic Excel implementation
2.7	Moderate	Sophisticated Excel implementation
4.0	High	Tedious and difficult to implement in Excel™, or best implemented outside Excel™ (specialist software required)
Score (Note 1)	Robustness rating	Descriptors
1.3	Yes	Works in different flow conditions with little to no alteration; not (or not very) sensitive to parameter settings or data errors (e.g., zeroes, outliers)
2.7	Partially	Works in different flow conditions with minor alteration; somewhat sensitive to parameter settings or data errors (e.g., zeroes, outliers)
4.0	No	Works in different flow conditions with major alteration; somewhat sensitive to parameter settings and requires data cleaning (removal of outliers)

Note 1: Criteria with three descriptor ratings have been adjusted to a numerical four-point score.

Table 9: MCA Weightings

Criteria	**Weightings**
Performance	50%
Data requirements	25%
Mathematical complexity	10%
Parameter complexity	5%
Robustness	10%

Table 10: Summary of Methods & MCA Analysis

Method	ADWF Estimate (ML/d)		Performance	Data Requirements	Mathematical Complexity	Parameter Complexity	Robust	Score[1]
	OSSTP	BVSTP						
20th Percentile	1.25	1.19	Excellent	Low	Low	1	Partially	**1.45**
50th Percentile	1.39	1.46	Moderate	Low	Low	1	Partially	**2.45**
Rainfall-based	1.35	1.32	Good	Moderate	Moderate	3	Yes	**2.25**
Histogram /Mode	1.30	1.16	Excellent	Low	Low	0	Partially	**1.43**
API Based[2]	1.31	1.22	Excellent	Moderate	Moderate	2	Yes	**1.73**
Ratio of Moving Averages (Arithmetic)[3]	1.30	1.19	Excellent	Low	Moderate	3	Yes	**1.50**
Ratio of Moving Averages (Arithmetic + Step-change)[4]	1.29	1.16	Excellent	Low	Moderate	5	Yes	**1.55**
Ratio of Moving Averages (EWMA)[5]	1.31	1.18	Excellent	Low	Moderate	4	Yes	**1.53**
Ratio of Moving Averages (EWMA + Step-change)[6]	1.30	1.16	Excellent	Low	Moderate	5	Yes	**1.55**
K-means Clustering[7]	1.53	1.68	Poor	Low	High	0	No	**3.33**
Diurnal Smoothing[8]	N/A	1.15	Excellent	High	High	2	Yes	**2.11**
Kernel Density Estimation[9]	1.30	1.22	Excellent	Very Low	High	1	Partially	**1.47**

[1] A lower score is better
[2] Decay Constant = 0.9, API Limit = 10, Median.
[3] Long term days = 28, Short term days = 7, Upper Bound = 1.025, Median.
[4] Long term days = 28, Short term days = 7, Upper Bound = 1.025, Step-change limit = 0.025, Median.
[5] Long term days = 28, Short term days = 7, Upper Bound = 1.025, Decay constant = 0.9, Median.
[6] Long term days = 28, Short term days = 7, Upper Bound = 1.025, Decay constant = 0.9, Step-change limit = 0.025, Median.
[7] K = 2
[8] Bandwidth = 2 hours, Outlier Factor = 5 relative to Median
[9] Bandwidth = Silverman's rule of thumb

Based on the data analysis and MCA undertaken for this study, the following observations are made regarding the various methods for ADWF estimation:

- The Percentile method can estimate ADWF very well, provided the appropriate percentile is chosen and the data has been filtered for errors and outliers. However, the appropriate percentile might vary depending on location and a degree of user input is required for data checking since the calculated percentile is somewhat sensitive to potential data errors (e.g., zeroes and outliers).
- The Histogram/ Mode method can estimate ADWF very well and is easy to perform. This method might require some degree of interpretation depending on location (if non-dry weather flows are very common). The Histogram is also useful in that it gives the user an opportunity to visualise the whole dataset, including outliers and zeroes.
- Rainfall-based methods can estimate ADWF very well with reasonably large datasets, provided strict rainfall thresholds are set. However, these methods require both daily flow data and daily rainfall data. The definition of appropriate rainfall thresholds might vary, depending on location, and proximity of the plant to a weather station with reliable rainfall records might be a constraint.
- The Ratio of Moving Averages method can estimate ADWF very well and does not require rainfall data. However, it requires a degree of user input for data manipulation and calibration around selection of upper and lower ratio bounds. Further notes for this method are:
  - Applying a step-change limit significantly improves the performance of the ratio of moving averages method when using mean (average) flow to calculate ADWF; whereas
  - Using the median flow instead of mean flow to calculate ADWF improves robustness of the ratio of moving averages method, largely eliminates the need for a step-change limit and reduces sensitivity to outlier data.
- K-means clustering can be effective for multi-modal flowrates. However, this method is not useful if flowrates are relatively uniformly distributed. This makes the method potentially sensitive to the location and nature of the sewer catchments.
- Diurnal Profile Smoothing can produce excellent estimates of ADWF, but it is much more mathematically complex and ‘data heavy’ since it requires short time interval flow data (hours or minutes) and outlier algorithms.
- Kernel Density Estimation can improve upon the Histogram/ Mode approaches where the number of data points is limited, or data is skewed. In most other respects, it does not offer significant advantages over simpler methods, given the disadvantage that it is mathematically complex.
- Several of the methods tested produce similar estimates of ADWF that can be considered reliable. Selection of the most appropriate method will likely depend on a balance of data requirements, mathematical simplicity (or complexity), and robustness (i.e., the extent to which user input/ interpretation required).

CONCLUSIONS

A review of existing industry ‘best practice’ methods for estimating ADWF at sewage treatment plants revealed that the current methods fall into three groups: Rainfall-based, EP-based, and Basic Statistical. Several potential Novel methods were also identified as a fourth group. Of those groups, Rainfall-based and Basic Statistical were the more common methods currently used. EP-based methods were not compared in this study, as they were considered to move the uncertainty from flow to population data, rather than attempting to directly calculate ADWF.

Based on this study, on balance of simplicity and performance against more complex methods, we recommend either or both of the following methods for routine estimation of ADWF at sewage treatment plants:

- Histogram/ Mode method - Plot histogram and/or calculate mode for datasets that are reasonably large (indicatively at least one year of daily STP total flow readings). Some user discretionary checking of data (at least from visual checks of histogram plots of raw data) is recommended to identify false zeroes and potential outliers, from which data filtering or outlier tests might be required.
- Percentile method: Adopt 20th percentile as a default, but with user discretion to check data (at least from visual checks of histogram plots of raw data, or using outlier tests), followed by minor sensitivity checks (e.g., test 20th, 30th and 50th percentiles).

For larger and more complex assignments (e.g., design projects, planning studies), it is recommended that one or more of the alternative high-performing methods described in this paper (e.g., Ratio of moving averages; Kernel Density Estimation) be employed for ADWF checks. Relatively large datasets (at least one year of daily flow totals) should be used and the results compared against the estimates from simpler methods (e.g., Histogram/ Mode and/ or Percentile methods).

ACKNOWLEDGEMENTS

We gratefully acknowledge the assistance of Anjila Finan (Byron Shire Council) with flow data collation and transmittal.

REFERENCES

Ali, S., Ghosh, N. C., & Singh, R. (2010, Mar 24). Rainfall-runoff simulation using normalized antecedent precipitation index. Hydrological Science Journal - Journal Des Sciences Hydrologiques, 55(2), 226-274.

City of Winnipeg. (2020). Wastewater Flow Estimation and Servicing Guidelines. Retrieved from Winnipeg Water and Waste Department, Winnipeg Manitoba: https://winnipeg.ca/waterandwaste/dept/wastewaterFlow.stm

Elias-Maxil, J. A., Hoek, J. P., Hofman, J., & Rietveld, L. (2014). A bottom-up approach to estimate dry weather flow in minor sewer networks. Water Science & Technology. 69.5, pp. 1059-1066. IWA Publishing.

GHD. (2017). Rainfall Analysis for Mullumbimby and Brunswick Heads WWTPs.

Kohler, M., & Linsley, R. (1951). Predicting The Runoff From Storm Rainfall. Washington: National Oceanic & Atmospheric Administration U.S. Dept. of Commerce.

Murray, N., Gabbett, T., & Townshend, A. (2017). Calculating acute:chronic workload ratios using exponentially weighted moving averages provides a more sensitive indicator of injury likelihood than rolling averages. British Journal of Sports Medicine, 51, 749-754.

NSWEPA. (2020). Environment Protection License 1802 for Kincumber Sewage Treatment System, Central Coast Council. Retrieved from NSW Environmental Protection Agency, Sydney: https://apps.epa.nsw.gov.au/prpoeoapp/

QLDEPA. (2019). Environmental Authority EPPR00887713 for Cairns Regional Council. Retrieved from Queensland Department of Environment and Science: https://apps.des.qld.gov.au/env-authorities/map/

QLDEPA. (2020). Environmental Authority EPPR00521513 for Queensland Urban Utilities. Retrieved from Queensland Department of Environment and Science: https://apps.des.qld.gov.au/env-authorities/map/

Queensland Water Supply Regulator, Water Supply and Sewerage Services, Department of Energy and Water Supply. (2014). Planning Guidelines for Water Supply and Sewerage. Retrieved from Queensland Government, Brisbane: https://www.dews.qld.gov.au/__data/assets/pdf_file/0016/80053/water-sewerage-planning-guidelines.pdf

UK Gov Environment Agency. (2018). Guidance Calculating dry weather flow (DWF) at waste water treatment works. Retrieved from United Kingdom Government, London: https://www.gov.uk/government/publications/calculating-dry-weather-flow-dwf-at-waste-water-treatment-works/calculating-dry-weather-flow-dwf-at-waste-water-treatment-works

USEPA. (2014). Guide for Estimating Infiltration and Inflow. Retrieved from United States Environmental Protection Agency, Washington DC: https://www3.epa.gov/region1/sso/pdfs/Guide4EstimatingInfiltrationInflow.pdf

Wikipedia. (2020a, July 21). Kernel density estimation. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/Kernel_density_estimation

Wikipedia. (2020b, Jul 21). Kernel smoother. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/File:Gaussian_Kernel_Regression.png

Wikipedia. (2020c, July 21). K-means clustering. Retrieved from Wikipedia: https://en.wikipedia.org/wiki/K-means_clustering

WSSA. (2014). Gravity Sewerage Code of Australia WSA Third Edition. Docklands, Victoria: Water Services Association of Australia Limited.

SUPPLEMENTARY INFORMATION

Table 11: Summary of Existing ADWF Estimation Methods

Method ID	Title	Definition	Source
1	QLD EPA Licenses: (North QLD – EPPR00887713)	A day in which no rain falls within the catchment of the sewage treatment plant from the commencement of measurement for that day. The term also excludes days during which measurement is made that occur within three (3) days following cumulative rainfall of 100 mm over the three (3) preceding days.	(QLDEPA, 2019)
2	QLD EPA Licenses: (SEQLD – EPPR00521513)	A day which less than 1 mm of rainfall is recorded at any rainfall measuring station recognised by the Commonwealth Bureau of Meteorology within the sewered are connected to the sewage treatment plant, or if no such measuring station exists, at the nearest such station to the sewage treatment plant. The term also excludes days during which recorded rainfall over the four (4) preceding days exceeds a cumulative rainfall 50 mm.	(QLDEPA, 2020)
4	NSW-EPA Licenses License 1802	For small plants, often no definition is given. For some larger plants, ADWF is defined as: *"…the average flow at a point calculated or measured over a 24 hour period in dry weather."* Where dry weather is defined as: *"Dry weather occurs when less than 10 millimetres of rainfall has been measured at a rain gauge in the catchment of the sewage treatment system during a 24 hour period (where there is no rain gauge in the catchment, at the rain gauge closest to the centre of the catchment)."*	(NSWEPA, 2020)
5	QLD Gov - Planning Guidelines for Water Supply and Sewerage (2014)	Planners should determine ADWF, PDWF, and PWWF based on: 1. Actual system performance 2. The WSAA Sewerage Code (see below) or 3. The historical Queensland approach, where typically: $ADWF\ is\ determined\ from\ first\ principles$[10] $PDWF = C_2 \times ADWF\ where\ C_2 = 4.7 \times EP^{-0.105}$ $PWWF = (5 \times ADWF)\ or\ (C_1 \times ADWF), whichever\ is\ the\ larger$ $C_1 = 15 \times EP^{-0.1587}\ (note: the\ minimum\ value\ for\ C_1 = 3.5)$	(Queensland Water Supply Regulator, Water Supply and Sewerage Services, Department of Energy and Water Supply, 2014).
6	WSAA Sewerage Code V.3.1	$ADWF = 0.0021 \times EP$ where ADWF is in L/s. Or the ML/d equivalent: $ADWF = 0.0001814 \times EP$.	(WSSA, 2014)

[10] Determine unit flow (L/EP) at ADWF for detached residential development based on internal water consumption and/or bulk metering of a residential catchment, and occupancy ratio. Determine total EP and total ADWF for detached residential development or land use. Determine unit flow (L/EP) at ADWF for all other (excluding detached residential) land use categories and/ or development/ customer categories. For all categories determine total EP and ADWF. Alternatively, where possible, determine total ADWF from treatment plant and catchment metering. Calibrate ADWF calculated from treatment plant or catchment metering against the calculated ADWF based on L/EP based land use categories. (Queensland Water Supply Regulator, Water Supply and Sewerage Services, Department of Energy and Water Supply, 2014).

Method ID	Title	Definition	Source
7	Previous planning work for Brunswick Heads and Mullumbimby	A day which receives <1 mm or <100 mm including the thirteen (13) preceding days.	(GHD, 2017)
8	US EPA: Average Dry Weather Flow	Flow during a period of extended dry weather (7 to 14 days) and seasonally high groundwater. Flow includes sanitary flow and infiltration, and excludes significant industrial and commercial flows (assumes no inflow during dry weather conditions)	(USEPA, 2014)
9	CA - Winnipeg Water and Waste Department	ADWF: $Number\ of\ Dewlling\ Units\ \times Number\ of\ People\ Per\ Unit \times 270\ L/capita/day$ Daily wastewater generation $= 270\ L/capita/day$ Single Family Dwelling: $Population/dwelling: 3.05$ $Dewlling/ha: 12.29$ Multi-Family Dwelling $Population/dwelling: 2.30$ $Dewlling/ha: 74.13$	(City of Winnipeg, 2020)
10	UK Gov: DWF Formula	$DWF = PG + I_{DWF} + E$ Where: $DWF = total\ dry\ weather\ flow\ (L/d)$ $P = catchment\ population\ (number)$ $G = per\ capita\ domestic\ flow\ (L/capita/d)$ $I_{DWF} = dry\ weather\ infiltration\ (L/d)$ $E = trade\ effluent\ flow\ (L/d)$	(UK Gov Environment Agency, 2018)
11	UK Gov: Nonparametric 20th percentile	The nonparametric 20-percentile value of a time series of measured total daily volume (TDV) data provides a good estimate of DWF. The 20-percentile number is that value exceeded by 80% of the recorded daily values. It is also known as the Q80.	(UK Gov Environment Agency, 2018)
12	Bottom-up approach (novel)	Bottom-up approach to estimate ADWF based on discharge inside households. Stochastic modelling based on flowrates and frequency of use (usage bursts). For relatively small systems where flowrate may be difficult to measure due to low flow.	(Elias-Maxil, et al., 2014))
13	Other novel methods	Ratio of short and long-term moving averages adopted from use in sports and human fitness.	(Murray, et al., 2017)

Table 12: Summary of ADWF Estimates Using Rainfall Method for Two STPs

OSSTP flow, Mullumbimby rainfall		ADWF estimates for different Average rainfall thresholds						Units
ADWF (ML/d): adopt 1.35	**Avg. rainfall (mm) incl. preceding days**	0	1	2	3	4	5	mm
No. preceding days	0	1.464	1.466	1.466	1.466	1.466	1.466	ML/d
	6	1.352	**1.354**	1.378	1.386	1.391	1.395	ML/d
	13	1.305	1.329	1.358	1.367	1.369	1.375	ML/d
	20	1.261	1.291	1.340	1.352	1.357	1.364	ML/d
	27	1.243	1.290	1.326	1.340	1.349	1.357	ML/d
BVSTP flow, Mullumbimby rainfall		**ADWF estimates for different Average rainfall thresholds**						
ADWF (ML/d): adopt 1.32	**No. of preceding days**	0	1	2	3	4	5	mm
No. preceding days	0	1.611	1.619	1.619	1.619	1.619	1.619	ML/d
	6	1.290	**1.319**	1.343	1.374	1.388	1.410	ML/d
	13	1.221	1.245	1.280	1.306	1.337	1.365	ML/d
	20	1.207	1.202	1.247	1.284	1.309	1.342	ML/d
	27	1.203	1.177	1.224	1.255	1.297	1.350	ML/d
OSSTP flow, Brunswick Heads rainfall		**ADWF estimates for different Average rainfall thresholds**						
ADWF (ML/d): adopt 1.37	**No. of preceding days**	0	1	2	3	4	5	mm
No. preceding days	0	1.462	1.468	1.468	1.468	1.468	1.468	ML/d
	6	1.376	**1.367**	1.388	1.392	1.398	1.402	ML/d
	13	1.370	1.347	1.366	1.377	1.380	1.384	ML/d
	20	1.403	1.314	1.350	1.365	1.367	1.373	ML/d
	27	1.436	1.318	1.339	1.355	1.360	1.366	ML/d
BVSTP flow, Brunswick Heads rainfall		**ADWF estimates for different Average rainfall thresholds**						
ADWF (ML/d): adopt 1.36	**No. of preceding days**	0	1	2	3	4	5	mm
No. preceding days	0	1.608	1.633	1.633	1.633	1.633	1.633	ML/d
	6	1.377	**1.357**	1.372	1.395	1.413	1.430	ML/d
	13	1.443	1.307	1.323	1.352	1.387	1.413	ML/d
	20	1.616	1.289	1.293	1.342	1.365	1.390	ML/d
	27	1.737	1.287	1.290	1.329	1.355	1.385	ML/d

Highlighted cells in table show adopted ADWF, based on 1 mm average rainfall over 7 days (i.e., any given day and 6 preceding days).

THE AUTHORS

Dr David de Haas

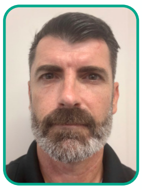
Dr David de Haas is a Senior Technical Director at GHD Pty Ltd with specialist skills in wastewater treatment systems. He has thirty-five years' experience in municipal water and wastewater treatment, covering research and development, planning, process design and operation. He has worked in wastewater-related consultancy in Australia for more than twenty years.

Stuart Ng

Stuart Ng is a Process Engineer at GHD Pty Ltd with skills in modelling and optimisation of wastewater treatment systems. He has 2 years' experience in process engineering, including hands-on operation and optimisation at a brewery waste treatment plant.

Nick Dahl

Nick Dahl is a Civil Engineer at GHD Pty Ltd with five years' experience, specialising in modelling water, wastewater and sewer systems, including the effects of rainfall. His work includes the use of InfoWorks ICM (integrated with ArcGIS), supplemented with the development of a range of analytics tools using MATLAB.



Dean Baulch

Dean Baulch is a Principal Engineer (Systems Planning, Utilities) at Byron Shire Council, a role he has filled for twelve years. He has thirty years of experience in the water industry, including planning, design, management and operation of urban water and wastewater systems.