University of Sheffield Experimental Manhole Traces and CRTDs Dataset Read Me

Dr Fred Sonnenwald

17th December 2020

1 Introduction

This dataset describes experimental solute traces (upstream and downstream temporal concentration profiles) recorded in model surcharged manholes at the University of Sheffield from 1993–2007. Deconvolved cumulative residence time distributions (CRTDs) for each trace are also provided with this dataset. It accompanies the journal article submission entitled "Predicting manhole mixing using a compartmental model". This dataset was collected by Dr Robert O'Brien (EPSRC grant GR/H43366/01), Dr Peter Dennis, Dr Chanwit Saiyudthong, and Dr Shing-Tak Douglas Lau. Dr Fred Sonnenwald performed the deconvolution to obtain the CRTDs and uploaded this archive under EPSRC grant EP/P012027/1.

Please visit http://mixing.group.shef.ac.uk for more information.

2 File naming and data format

This dataset consists of six `ZIP` file archives, one `XLS` spreadsheet, and this `PDF` Read Me document. The first five `ZIP` files contain solute traces for each diameter of manhole studied (see Section 3 for more information on the experimental setup and manhole configurations). The last `ZIP` file contains the deconvolved CRTDs. Each solute trace `ZIP` file contains a number of folders, one for each flow rate studied, e.g. `1.00 lps flow rate`, or outlet angle, e.g. `30 degree outlet`. The latter contains flow rate folders. The flow rate folders contain `CSV` data files of the concentration profiles (traces) for the straight-through manhole configuration. The flow rate folders may also contain additional subfolders for the configuration variants in the case of a stepped outlet, e.g. `132 mm step` or benching, e.g. `benched`, which contain additional `CSV` data files. In this way, each data file can be associated with an experimental configuration. The `ZIP` file containing the deconvolved CRTDs is organised in the same way. The `XLS` spreadsheet contains a summary of all experimental configurations and associated temporal concentration profiles.

2.1 Temporal concentration profiles

Each folder containing `CSV` files contains the data files for all surcharge depths and repeat injections for that manhole configuration. Note that both the number of surcharge depths and number of repeats vary between configurations. There are 4,969 traces in total.

In addition to the organisation by directory structure, each file name provides configuration identification (CID) and file identification (FID) numbers to use in conjunction with the XLS spreadsheet to uniquely identify each data file with each experimental configuration. For example, take the file `388 mm diameter.zip/90 degree outlet/8.00 lps flow rate/benched/CID 0854 FID 3825 - 0215 mm surcharge repeat 1.csv`. The directory structure says the data were collected from a benched 388 mm manhole with a 90° outlet at 8 ls^{-1}. The filename says the data were collected at a 215 mm surcharge depth and that the data are from repeat injection 1.

The `CID 0854 FID 3825 - 0215 mm surcharge repeat 1.csv` filename also provides a configuration identification number of 854 and file identification number of 3825. By searching for the CID number in the XLS spreadsheet, we can see the configuration details below match the file name and directory structure. The XLS spreadsheet also contains the distance between the manhole centre and the fluorometers, the number of repeat injections, and the file names and paths for the other repeat injections.

Configuration ID Number	Flow Rate (l/s)	Surcharge Depth (mm)	Manhole Internal Diameter (mm)	Pipe Internal Diameter (mm)	Angle (degrees)	Step Height (mm)	UNBENCHED or BENCHED	Centreline to Fluorometer (mm)	...
...									
853	8.00	188	388	88	30	0	BENCHED	1350	...
854	8.00	215	388	88	30	0	BENCHED	1350	...
855	8.00	241	388	88	30	0	BENCHED	1350	...
...									

...	Reference	Repeats	Starting ID Number	File 1	File 2	File 3
...						
...	Saiyudthong (2004)	3	3822	...repeat 1.csv	...repeat 2.csv	...repeat 3.csv
...	Saiyudthong (2004)	3	3825	...repeat 1.csv	...repeat 2.csv	...repeat 3.csv
...	Saiyudthong (2004)	3	3828	...repeat 1.csv	...repeat 2.csv	...repeat 3.csv
...						

Each CSV data file contains three columns: time (in seconds), upstream concentration, and downstream concentration. The concentration data are the uncalibrated sensor reading in volts when possible (see Section 4 for more information on calibration). To illustrate, the first part of `CID 0854 FID 3825 - 0215 mm surcharge repeat 1.csv` contains:

```
0.484848485,        0.156555176,        0.260620117
0.515151515,        0.154724121,        0.260314941
0.545454545,        0.16418457,         0.272521973
0.575757576,        0.15838623,         0.270080566
...,                ...,                ...,
```

The data from `CID 0854 FID 3825 - 0215 mm surcharge repeat 1.csv` are plotted in Figure 1.

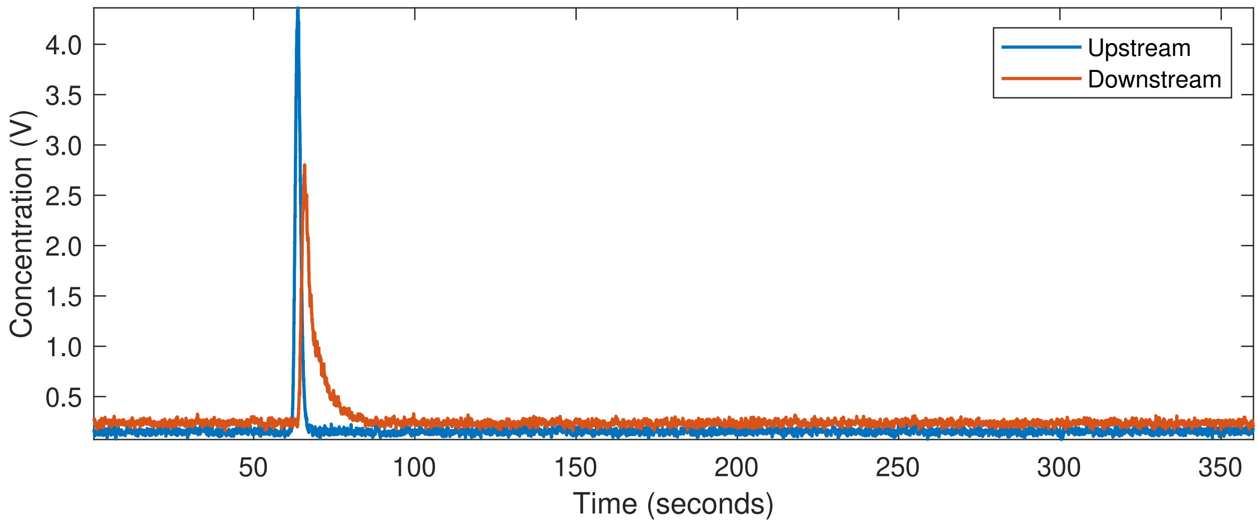


Figure 1: Example plot of the uncalibrated benched 388 mm manhole 8 ls^{-1} 215 mm surcharge 90° outlet configuration (CID 854, FID 3825)

2.2 Deconvolved CRTDs

Each solute trace has a deconvolved cumulative residence time distribution (CRTD) derived from the upstream and downstream temporal concentration profiles. The CRTD `CSV` files are organised and named the same way as the temporal concentration profiles with `CRTD` added to the file name before the file extension. Each `CSV` file contains two columns: time (in seconds) and deconvolved CRTD (see Section 5 for more information on deconvolution). To illustrate, the first part of `CID 0854 FID 3825 - 0215 mm surcharge repeat 1 CRTD.csv` contains:

```
0,          0.000187964732
0.1,        0.000376631237
0.2,        0.000566250937
0.3,        0.000757075254
...,        ...
```

The data from `CID 0854 FID 3825 - 0215 mm surcharge repeat 1 CRTD.csv` are plotted in Figure 2.

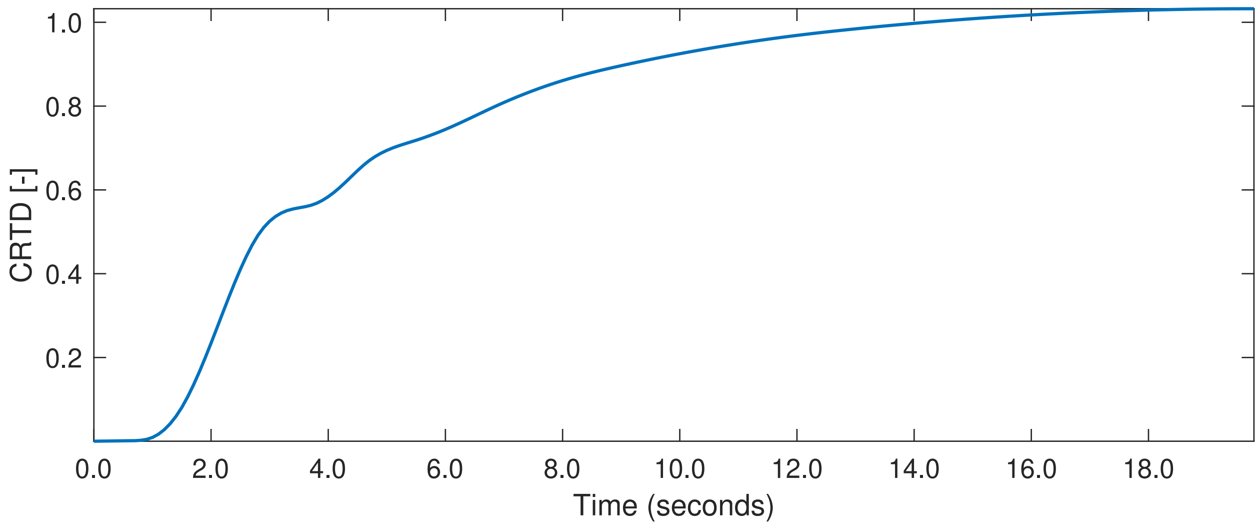


Figure 2: Example plot of the CRTD for repeat inject 1 of the benched 388 mm manhole 8 ls^{-1} 215 mm surcharge 90° outlet configuration (CID 854, FID 3825)

3 Experimental setup

O'Brien (2000), Dennis (2000), Saiyudthong (2004), and Guymer *et al.* (2005) conducted solute tracing in a series of model manholes with diameters of 388, 500, 600, and 800 mm and an 88 mm diameter inlet pipe. The inlet and outlet pipe were fitted with Turner Designs Series 10 fluorometers (Turner Designs, San Jose, California) located 1350 mm from the manhole centre and fitted with a custom large pipe adapter device to measure dye concentrations. Lau (2008) conducted solute tracing in a model manhole at a different scale with a 218 mm diameter and a 24 mm diameter inlet pipe. The inlet and outlet pipe were fitted with Series 10 Turner Design fluorometers (without adapter bracket) located 368 mm from the manhole centre. Surcharge level was measured by a water level follower and flow rate by venturi meter. Surcharge level was controlled by a downstream storage tank with variable height weir. Rhodamine WT (O'Brien, 2000; Dennis, 2000; Saiyudthong, 2004; Guymer *et al.*, 2005) or 6G (Lau, 2008) dye was injected into the pipe several metres upstream of the manhole. Between 2 and 8 repeat injections were carried out. Smooth pipes were used in the experimental work with a relative roughness of approximately 2×10^{-5} and friction factor of 0.020. A schematic of the experimental setup is shown in Figure 3 and photographs are shown in Figure 4. Additional details on the experimental setups are available in the relevant references.

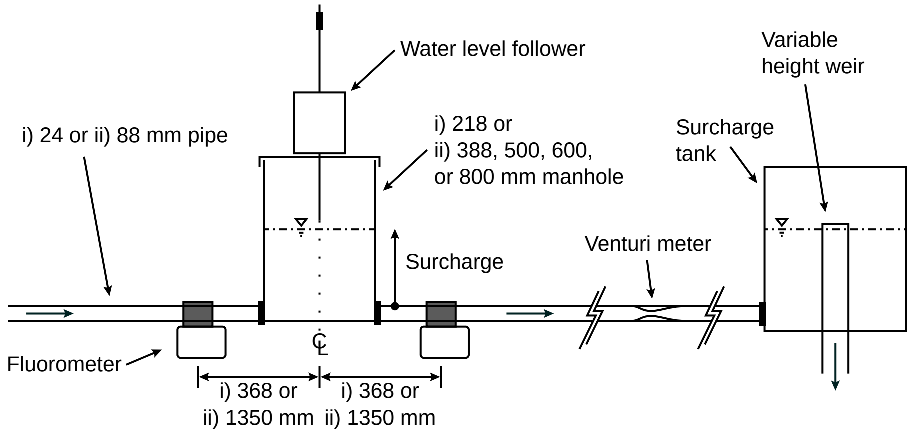
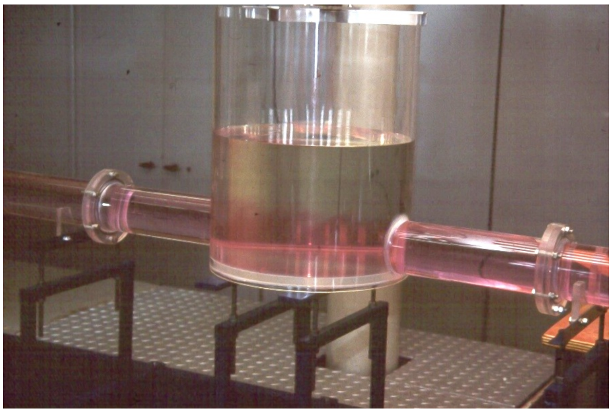


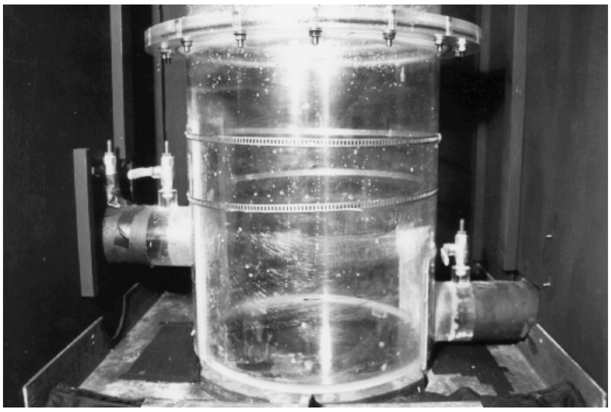
Figure 3: Experimental model manhole setup overview, after Lau (2008), experiments were conducted at either i) 24 mm or ii) 88 mm pipe diameter scale

The solute traces were conducted in a range of manhole configurations varying diameter, flow rate, surcharge depth, outlet angle, outlet step and benching. Surcharge depths, measured above the outlet soffit, varied between 0 and 1242 mm. Flow rate varied between 0.25 and 8.77 ls^{-1}. Outlet step height varied between 0 and 172 mm. Outlet angle varied between 0 and 90°. Some manhole configurations had 1:12 benching and a full pipe depth dry weather flow channel. The range of manhole configurations is given in Table 1. There are a total of 1,153 configurations, the full details of which are in the XLS spreadsheet.

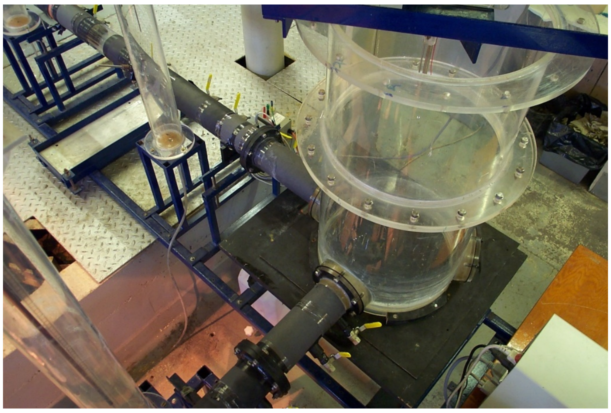
(a) 388 mm straight-through manhole, CID 1–96 (O'Brien, 2000)



(b) 388 mm straight-through manhole, CID 1–96, showing dye (O'Brien, 2000)



(c) 388 mm diameter, 132 mm stepped unbenched manhole, CID 489–555 (Dennis, 2000)



(d) 388 mm 90° angled unbenched manhole, CID 983–1049 (Saiyudthong, 2004)

Figure 4: Experimental model manholes

Table 1: Manhole configurations

CID	Manhole Diameter D_m (mm)	Pipe Diameter D_p (mm)	$\frac{D_m}{D_p}$	Flow Rate ($\mathrm{l\,s^{-1}}$)	Surcharge Depth S (mm)	$\frac{S}{D_m}$	Outlet Angle	Step Height B (mm)	$\frac{B}{D_p}$	Reference
1104–1153	218	24	9.1	0.25–0.5	7–100	0.04–0.46	0°	0	0	Lau (2008)
1–96, 266–663, 664–1049	388*	88	4.4	1–8	1–1241	0.0–3.2	0°, 30°, 60°, 90°	0, 44, 88, 132, 176	0, 0.5, 1.0, 1.5, 2	O'Brien (2000); Dennis (2000); Saiyudthong (2004)
97–156	500	88	5.7	0.97–7.45	7–268	0.02–0.54	0°	0	0	O'Brien (2000)
157–240	600	88	6.8	0.87–8.72	27–308	0.04–0.51	0°	0	0	O'Brien (2000)
241–265, 1050–1103	800	88	9.1	0.88–8.77	27–332	0.03–0.41	0°	0	0	O'Brien (2000); Guymer *et al.* (2005)

*Only some stepped and angled 388 mm manholes experiments had benching

4 Calibration

Table 2 provides the calibration functions obtained with the original experimental data. The calibration functions are given for varying configurations, giving concentration in parts-per-million (PPM) as a function of voltage reading (V). Note, data from Guymer *et al.* (2005) is supplied already calibrated as the raw data files could not be located.

Table 2: Calibration functions

	CID	Range	Function	Reference
Upstream Downstream	1–265	All	$PPM = 5.555 \times 10^{-1}V - 9.44 \times 10^{-2}$ $PPM = 1.722 \times 10^{-1}V + 1.21 \times 10^{-2}$	O'Brien (2000)
	266–309	$B/D_p = 0.0$	$PPM = 2.553 \times 10^{-1}V - 3.07 \times 10^{-2}$ $PPM = 3.195 \times 10^{-1}V - 3.37 \times 10^{-2}$	Dennis (2000)
	310–400	$B/D_p = 0.5$	$PPM = 3.124 \times 10^{-1}V - 3.75 \times 10^{-2}$ $PPM = 1.826 \times 10^{-1}V - 2.04 \times 10^{-2}$	
	401–488	$B/D_p = 1.0$	$PPM = 2.845 \times 10^{-1}V - 1.47 \times 10^{-2}$ $PPM = 1.710 \times 10^{-1}V - 2.87 \times 10^{-2}$	
	489–583	$B/D_p = 1.5$	$PPM = 3.255 \times 10^{-1}V - 4.52 \times 10^{-2}$ $PPM = 1.697 \times 10^{-1}V - 2.48 \times 10^{-2}$	
	584–663	$B/D_p = 2.0$	$PPM = 3.362 \times 10^{-1}V - 3.81 \times 10^{-2}$ $PPM = 1.753 \times 10^{-1}V - 2.48 \times 10^{-2}$	
	664–727, 862–918	30°	$PPM = 3.519 \times 10^{-1}V - 4.84 \times 10^{-2}$ $PPM = 3.517 \times 10^{-1}V - 1.49 \times 10^{-1}$	Saiyudthong (2004)
	728–797	60° Benched	$PPM = 2.619 \times 10^{-1}V - 2.49 \times 10^{-2}$ $PPM = 2.057 \times 10^{-1}V - 1.49 \times 10^{-2}$	
	919–982	60° Unbenched	$PPM = 2.188 \times 10^{-1}V + 6.41 \times 10^{-3}$ $PPM = 1.643 \times 10^{-1}V - 4.96 \times 10^{-2}$	
	798–861, 983–1049	90°	$PPM = 2.514 \times 10^{-1}V - 4.73 \times 10^{-2}$ $PPM = 1.917 \times 10^{-1}V - 2.40 \times 10^{-2}$	
	1050–1103	All	Data already calibrated	Guymer *et al.* (2005)
	1104–1153	All	$PPM = 4.373 \times 10^{-1}V - 3.04 \times 10^{-2}$ $PPM = 4.578 \times 10^{-1}V - 1.76 \times 10^{-2}$	Lau (2008)

5 Deconvolution

This dataset provides a cumulative residence time distribution (CRTD) for each solute trace. In general the downstream concentration profile $C(x_2,t)$ is related to the upstream concentration profile $C(x_1,t)$ through convolution

$$C(x_2,t) = \int_{-\infty}^{\infty} E(\tau)C(x_1,t-\tau)d\tau \tag{1}$$

where E is the residence time distribution (RTD) and τ is an integration variable. The CRTD is the cumulative sum of the RTD. The inverse operation of Equation 1, using known upstream and downstream data to obtain the RTD, is called deconvolution. Deconvolution as described in Stovin *et al.* (2010), Sonnenwald *et al.* (2014), and Sonnenwald *et al.* (2015) has been used, with some minor modifications. The modifications are described below.

5.1 Deconvolution method modifications

The first modification is on the sampling point method used by the deconvolution routine. Sonnenwald *et al.* (2014) describes a method where the sample points were placed where the slope of the FFT initial RTD guess was greatest. This approach has been modified for the deconvolution used here. The first sample point after $t = 0$ is now located at 70% of the difference in first arrival time (defined as 0.2% of the peak of the concentration profiles). The last sample point is now located at $t_{99.8}$, the time at which the FFT initial CRTD guess takes a value of 0.998. The sample points are placed between these two times as previously described by Sonnenwald *et al.* (2014), based on the slope of the FFT initial RTD guess.

The change to the sample point spacing allows for the RTD/CRTD to be shorter than the recorded data, reducing the need for trimming when the time-series is long compared to the duration of the trace. The same number of sample points can therefore result in a higher resolution sub-sampling compared to the previous sample point spacing method and fewer sample points can be used to the same effect. Thus, the second modification involves the number of sample points used. When deconvolving the current dataset, deconvolution was initially carried out using only 15 sample points. If the R_t^2 goodness-of-fit between the recorded downstream concentration profile and the predicted downstream concentration profile (using the deconvolved RTD) was less than 0.98 the number of sample points was continuously increased by 5 and deconvolution repeated until the R_t^2 exceeded 0.98, the change in R_t^2 was less than 0.01, or 50 sample points were reached.

5.2 Data pre-processing

Initial pre-processing was carried out to apply the reported calibrations, filter noise, subtract background concentration, and trim the traces. A Butterworth filter was applied with a 1 Hz cut-off frequency to remove high-frequency sensor noise. Background subtraction was taken as the mean of the first and last 10 seconds of data. A copy of the trace was smoothed with a moving average and used to determine the start time of the trace as 1% of the smoothed peak upstream profile and end of the trace as 1% of the peak of the smoothed downstream profile to ensure a conservatively large cut-off.

Secondary data pre-processing consisted of applying a moving average and then re-sampling to 10 Hz. For data with time-steps greater than 10 Hz and less than 100 Hz a 50 time-step moving average was applied and data with a time-step greater than or equal to 100 Hz had a 500 time-step moving average applied. Additional manual background subtraction or trimming was applied to less than 1.5% of traces to raise R_t^2. Mean R_t^2 of the experimental downstream concentration profile compared to the downstream concentration profile predicted by convolving the upstream profile with the deconvolved RTD was 0.997 and the standard deviation was 0.004. Only 0.2% of RTDs had an R_t^2 value of less than 0.98. The RTDs were saved as CRTDs for plotting and interpretation of results.

References

Dennis, P. (2000). *Longitudinal Dispersion due to Surcharged Manholes*. PhD thesis, The University of Sheffield. https://etheses.whiterose.ac.uk/14802/

Guymer, I., Dennis, P., O'Brien, R., & Saiyudthong, C. (2005). Diameter and surcharge effects on solute transport across surcharged manholes. *Journal of Hydraulic Engineering*, 131(4), 312–321. https://doi.org/10.1061/(asce)0733-9429(2005)131:4(312)

Lau, S.-T. D. (2008). *Scaling Dispersion Processes in Surcharged Manholes*. PhD thesis, The University of Sheffield. https://etheses.whiterose.ac.uk/14935/

O'Brien, R. (2000). *Dispersion due to Surcharged Manholes*. PhD thesis, The University of Sheffield. https://find.shef.ac.uk/permalink/f/15enftp/44SFD_ALMA_DS21181058280001441

Saiyudthong, C. (2004). *Effect of Changes in Pipe Direction across Surcharged Manholes on Dispersion and Head Loss*. PhD thesis, The University of Sheffield. https://etheses.whiterose.ac.uk/14847/

Sonnenwald, F., Stovin, V., & Guymer, I. (2014). Configuring maximum entropy deconvolution for the identification of residence time distributions in solute transport applications. *Journal of Hydrologic Engineering*, 19(7), 1413–1421. https://doi.org/10.1061/(ASCE)HE.1943-5584.0000929

Sonnenwald, F., Stovin, V., & Guymer, I. (2015). Deconvolving smooth residence time distributions from raw solute transport data. *Journal of Hydrologic Engineering*, 20(11), 04015022. https://doi.org/10.1061/(ASCE)HE.1943-5584.0001190

Stovin, V. R., Guymer, I., Chappell, M. J., & Hattersley, J. G. (2010). The use of deconvolution techniques to identify the fundamental mixing characteristics of urban drainage structures. *Water Science and Technology*, 61(8), 2075–2081. https://doi.org/10.2166/wst.2010.134