



High resolution hydrometric and sewer system monitoring of an urban catchment with complex flood risk

Chris Sweetapple¹, Tegan Illingworth², Peter Melville-Shreeve¹

¹ Centre for Water Systems, College of Engineering, Mathematics and Physical Sciences, University of Exeter, North Park Road, Exeter, Devon EX4 4QF, United Kingdom

² East Sussex County Council, County Hall, St Anne's Crescent, Lewes, East Sussex, BN7 1UE, United Kingdom

Correspondence to: Chris Sweetapple (C.Sweetapple@exeter.ac.uk)

Abstract. Stormwater management systems are under increasing pressure from challenges such as urbanisation and climate change, and better understanding and management is essential to reduce flood risk, protect water quality and ensure the resilience of urban infrastructure. Traditionally, there has been a strong reliance on physics-based hydrological and hydraulic models; however, model-based approaches are subject to limitations and can be challenging to implement in many scenarios. Consequently, there is now growing interest in the use of data-driven methods to complement or augment traditional modelling approaches. However, the scarcity of openly available, integrated hydrometric and sewer system monitoring data poses a significant barrier to the development, validation and implementation of 'smarter' stormwater management systems, and continues to limit progress in data-driven approaches. To address this gap, this paper provides detailed, high density hydrometric and sewer level data collected from 497 new gauges across 78km² urban catchment in the United Kingdom, including from rainfall gauges, fluvial water level gauges, groundwater monitoring boreholes, road gullies and in-sewer level sensors. This data is supplemented by information on the sewer system design, fluvial network and catchment topology, and scripts for accessing, processing and visualising the data. It is intended that publishing this large-scale, high resolution urban hydrometric and sewer level dataset will support development of improved methods for understanding urban stormwater systems, advance data-driven modelling approaches, and enable the next generation of smart stormwater management strategies. Several potential research opportunities are discussed in detail, including understanding spatio-temporal variability in urban hydrological processes and responses; improving sensor network design; machine learning and artificial intelligence applications; characterising groundwater-sewer interactions; improving data-driven asset-management and decision making; and application in benchmarking and open science applications more broadly.

The dataset repository is available at <https://doi.org/10.5281/zenodo.20699634> (Sweetapple et al., 2026a) and example scripts at <https://doi.org/10.5281/zenodo.21390823> (Sweetapple et al., 2026b).

1 Introduction

Stormwater management systems are facing increasing pressure due to challenges such as urbanisation and climate change. Rapid growth in impervious surface areas is altering hydrological processes, reducing infiltration and evapotranspiration and



increasing runoff volumes and peak flows (Berndtsson, 2010; Fletcher et al., 2013), whilst many locations are expected to experience more frequent and more intense storm events due to the impact of climate change on precipitation patterns (Mallakpour and Villarini, 2017). As such, the frequency of pluvial flooding and sewer surcharges is expected to increase in many regions (Hauser et al., 2026; Hosseinzadehtalaei et al., 2021; Wartalska et al., 2025). Better understanding and management of urban stormwater systems is, therefore, essential in order to reduce flood risk, protect water quality, and ensure the resilience of urban infrastructure.

These systems are complex networks that integrate surface hydrology, groundwater processes, sewer hydraulics and operational control. Their behaviour is highly nonlinear and strongly influenced by characteristics such as spatial and temporal heterogeneity in rainfall, topography, land use, and infrastructure design and operation. Understanding and management of such systems has traditionally relied heavily on physics-based hydrological and hydraulic models such as SWMM, MIKE+ and InfoWorks ICM, with these having become essential tools for design, planning and flood risk assessment. However, model-based approaches are subject to limitations and can be challenging to implement in many scenarios – for instance, due to gaps in system knowledge hindering model setup, difficulties in representing real-world operational conditions, high computational demands preventing real-time simulations, limited availability of calibration data, availability of the required expertise, and licencing costs (e.g. Cea et al., 2025). Furthermore, uncertainties in drainage connectivity, infiltration processes and system operation can significantly affect model performance (Li et al., 2025b; Wang et al., 2019).

Consequently, as distributed sensor networks capable of providing high-frequency hydrometric observations in near real-time become more accessible due to advances in sensing technologies, telemetry and cloud computing, there is now growing interest in the use of data-driven methods such as machine learning, artificial intelligence (AI) and digital twins to complement or augment traditional modelling approaches.

Data-driven methods have the potential to capture complex system dynamics without the need to explicitly model processes such as infiltration, groundwater interactions or operational control, and studies to date have demonstrated their benefit for applications such as flood forecasting (Kuhaneswaran et al., 2025), real-time control (e.g. Mullapudi et al., 2020; Sharior et al., 2019) and anomaly detection (e.g. Qiu et al., 2024). However, despite their potential, transition towards ‘smarter’, data-driven stormwater management remains slow, with many (real or perceived) technical and socio-economic barriers related to the availability and reliability of technologies, technological and physical limitation, decision making, uncertainty, security, lack of confidence, resistance to change, expense, ownership, lack of knowledge and the need for community engagement remaining (Sweetapple et al., 2023). Monitoring of stormwater systems thus remains typically sparse, and availability of open, long-term, high spatio- and temporal-resolution observational datasets from urban drainage systems which could be used to address some of the identified barriers is very limited. As such, a lack of comprehensive datasets remains a fundamental barrier to the adoption of AI in stormwater management (Rahman et al., 2026).

There have been some recent publications that make significant progress towards addressing this gap: For instance, the Bellinge dataset (Nedergaard Pedersen et al., 2021) provides up to 10 years of sewer and rainfall observation data from a small (1.7 km²) urban drainage system in Denmark, along with information on the system characteristics and hydrodynamic models, whilst



65 the Urban Water Observatory dataset from Switzerland provides three years of high-resolution monitoring from 124 sensors capturing precipitation and meteorological data, wastewater hydraulics, and wastewater temperature from a 28.5 km sewer system. Both highlight the importance of providing accompanying metadata, system information and contextual geographical data to enable effective use of the monitoring datasets for novel applications, and also supply valuable asset data and urban drainage system models. However, the focus of both of these studies is on the sewer system, and there remains a need for
 70 comprehensive datasets that will facilitate study into coordinated stormwater management at a catchment level (e.g. addressing fluvial, pluvial and groundwater flooding in addition to sewer flooding) via inclusion of wider hydrometric data from interacting systems.

This scarcity of openly available, integrated hydrometric and sewer system monitoring data poses a significant barrier to the development, validation and implementation of ‘smarter’ stormwater management systems, and continues to limit progress in
 75 data-driven approaches. To address this gap, this paper provides detailed, high density hydrometric and sewer level data collected across an urban catchment in the Eastbourne and southern Wealden area of the United Kingdom. The study area covers approximately 78km² and includes a complex hydrological system comprising a predominantly combined sewer system, fluvial channels, storage lakes and ponds and a chalk catchment. Much of the catchment is very flat, which contributes to difficulties in modelling flows, and there has historically been a poor understanding of how water from different sources
 80 interact. However, with the area at risk from pluvial, fluvial, groundwater and coastal flooding, flood management is a complex problem and there is impetus to better understand and manage the system through data-driven approaches. Whilst an integrated 1D-2D hydraulic model (developed in InfoWorks) has provided improved understanding of how the catchment responds to a library of different scenarios, the model is complex and slow to run (taking around two weeks per simulation), and thus its applications are limited. High spatial and temporal resolution observations have, therefore, been collected from 497 new
 85 monitoring locations, including rainfall gauges, fluvial water level gauges, groundwater monitoring boreholes, road gullies and in-sewer level sensors, and this data is supplemented by information on the sewer system design, fluvial network and catchment topology extracted from the integrated model.

For context, prior to this study, data was publicly available via the Environment Agency from only four level gauges and one rain gauge in the 78km² catchment, and no sewer level data was available. The new monitoring density (average one sensor
 90 for every 0.16 km²) also contrasts significantly to the England average of one sensor for every 14 km², based on the approximately 9200 rainfall, river flow/level and groundwater level sensors accessible via the Environment Agency’s Hydrology Data Explorer (DEFRA, 2024).

This dataset was generated as part of the Blue Heart project (Blue Heart, 2025), which aims to transform understanding of the system and facilitate smarter stormwater management to improve flood risk management and resilience through environmental
 95 monitoring and digital infrastructure. To maximise the benefits of the data generated for this project, it is provided here for use by the wider research community. The hydrometric and sewer level observations are provided as curated time series and metadata, and these are enriched with geographical and topographical information describing the catchment, sewer system and monitoring infrastructure, and contextual environmental data. In addition, example scripts for accessing, processing and



visualising the data are provided. These resources aim to lower the barrier to entry for researchers seeking to explore the data or develop new analyses and modelling approaches. It is intended that publishing this large-scale, high resolution urban hydrometric and sewer level dataset will support development of improved methods for understanding urban stormwater systems, advance data-driven modelling approaches, and enable the next generation of smart stormwater management strategies.

2 Material and methods

2.1 Catchment description

The case study sensor network is located in the county of East Sussex, situated on the south coast of the United Kingdom (as shown in Figure 1a). The study area covers approximately 78 km² and includes the urban areas of Eastbourne and southern Wealden. It is bordered by the English Channel to the south and includes parts of the South Downs (a prominent chalk escarpment) to the west. It also includes part of the Pevensey Levels, which comprise of an extensive area of low lying wet grassland, much of which is reclaimed marshland and includes artificial drainage channels and embankments.

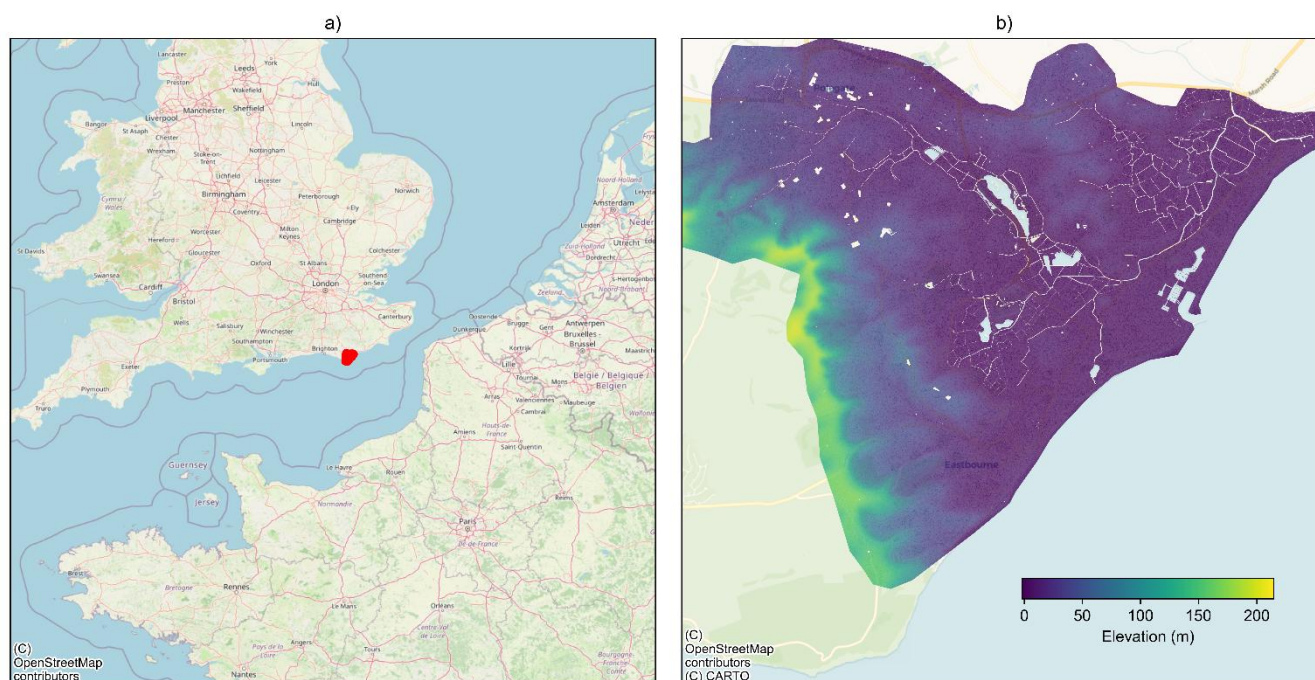


Figure 1: a) Location of case study catchment, shown in red; and b) Catchment elevation.

The catchment is predominantly low-lying, with over 50% below 15 mAOOD (metres above ordnance datum), 0.17 km² below 0 mAOOD, and a minimum elevation of -1.8mAOOD. For context, sea levels are typically in the range -1.8 – 3.7 mAOOD. Steep



115 slopes are found along the western edge, leading to higher elevations in the range 150-200 mAOD. Elevations across the catchment are shown in Figure 1b.

The region experiences a temperate maritime climate, influenced by its coastal proximity and influence of the North Atlantic. Long-term climate observations indicate mean annual rainfall of approximately 793 mm, which is distributed throughout the year but with more falling during the autumn and winter months than spring and summer (Met Office, 2026). Annual mean
 120 maximum temperatures are 14.6 °C (Met Office, 2026), with relatively mild winters and moderate summer temperatures compared with inland areas due to the maritime influence.

Land uses within the catchment include urban areas (encompassing residential, tourism-related, commercial and industrial development), as well as agricultural and semi-natural landscapes. The urban area of Eastbourne, which has a population of approximately 104,000, forms the dominant land use. Parts of the western catchment fall within the South Downs National
 125 Park, which contains protected chalk grassland ecosystems and areas managed for conservation and recreation. There are also several areas of parks, the largest of which is Shinewater Park.

With respect to geology, the north-east of the catchment contains bands of marine-origin sedimentary rocks (including Weald Clay Formation Mudstone; Lower Greensand Group Sandstone, siltstone and mudstone; and Gault Formation Mudstone), running north-west to south-east, whilst the south-west contains chalk bands of various types (including West Melbury Marly
 130 Chalk Formation, Zig Zag Chalk Formation, Holywell Nodular Chalk Formation, New Pit Chalk Formation, Lewes Nodular Chalk Formation and Lewes Nodular Chalk Formation). The superficial geology includes fluvial-origin alluvium clay, silt, sand and peat in low lying areas; a band of storm beach deposits gravel along coastline; and patches of head-clay silt, sand and gravel in the west (BGS, 2026). The chalk formations form a major aquifer with significant groundwater storage capacity, and allow rapid infiltration and groundwater movement.

135 This combination of permeable chalk geology, urban development and low-lying coastal drainage systems result in a complex environment of interacting groundwater dynamics, surface water processes and engineered drainage infrastructure.

2.2 Hydrology, stormwater systems and flood risk

2.2.1 Watercourses

A network of small watercourses and channels with a total length of over 85 km drain the catchment towards the coast, as
 140 shown in Figure 2. The hydrological behaviour is influenced by multiple artificial storage bodies (Section 2.2.2) and tidal boundary conditions affecting the outfalls (Section 2.2.3). There is also a pumping station on the Lottbridge Sewer that is operated by the Environment Agency, and multiple flow control structures in the network (note: several watercourses in the area are named '[...] Sewer'; these are part of the fluvial system and not the sewer network). Key control structures include:

- Fence Bridge Gates, which consist of three adjacent gates located at the end of the Langney Haven. These are
 145 maintained by the Environment Agency and operated according to 'summer' and 'winter' rules (based on upstream



and downstream levels) to manage water on the Pevensey Levels. An additional structure south of Fence Bridge Gates is operated similarly.

- Crumbles Gate, which is located at the inlet to Crumbles Pond (Section 2.2.2) and is operated manually. It should be open if the upstream level exceeds a specified threshold, but the strategy has not been possible to verify.

150 Further gates exist, but are typically operated manually and/or not in use (i.e. permanently shut or permanently open).

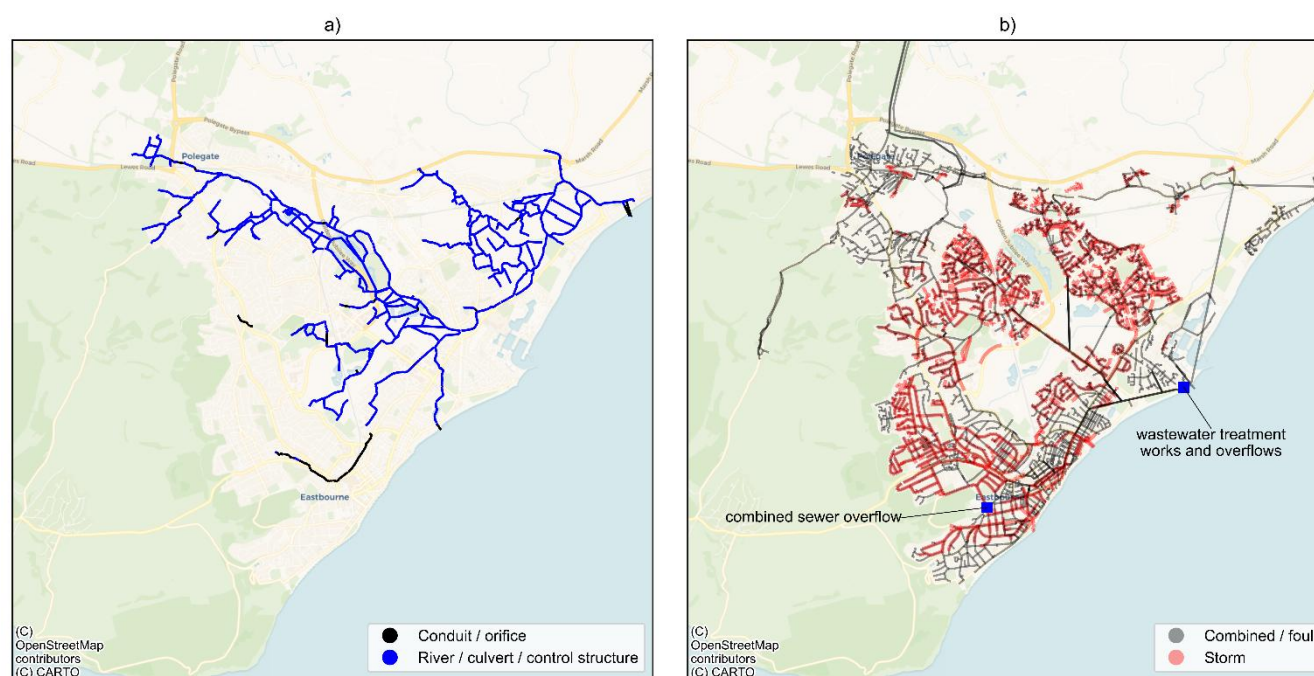


Figure 2: a) Fluvial / overland system; and b) stormwater, combined and foul sewer system (extracted from the hydraulic model of the case study area (discussed in Section 1)).

2.2.2 Storage

155 The case study area contains a series of four flood attenuation lakes (Shinewater Lake, Langney Lake, West Langney Lake and Hydneye Lake) that were constructed in the early 1990s to manage flood risk from new development. All have since suffered from heavy siltation, which has reduced their capacity. Management is made challenging by the presence of Bronze Age remains in Shinewater Lake, the presence of a Scheduled Ancient Monument, and designation as a Local Nature Reserve. There is also the Crumbles Pond – a large, artificial storage body located in Princes Park that is used for recreational purposes.

160 This receives controlled inflows from the Crumbles Sewer, and controls outflows to the English Channel via a weir.

2.2.3 Outfalls

There are two fluvial outfalls to the sea which convey flows via flap valves. One (the Crumbles outfall) is located in the main town, downstream of the Crumbles Pond; the other is 5.2 km away at Pevensey Bay and serves flows from Eastbourne, the



western Pevensey Levels and Hailsham. These outfalls, however, are vulnerable to tidal locking (where no discharge is possible during high tides), which can result in water accumulating upstream in the drainage system. There are historical examples of this causing flooding and significant damage when aligned with a storm event (for instance a 39.4mm twin-peaked storm in August 2015 that coincided with high tide led to flooding of the town centre and resulted in extensive damage to businesses and residential properties), and this issue is expected to worsen with sea level rises resulting from climate change.

2.2.4 Sewer system

Urban drainage infrastructure also plays a role in stormwater management in the urban Eastbourne area, with the town served by a partially combined (foul and storm) and partially separate sewer system (Figure 2b). This drains to a wastewater treatment works in the north east of the catchment (50.7847° N, 0.3264° E (WGS84)), which provides dry weather and storm flow capacities (throughput) of 303 l/s and 858 l/s respectively and serves a population equivalent of approximately 117,439. The wastewater treatment works provides storm management via storm overflows and emergency overflows and is connected to an 85m short-sea outfall that may be affected by tide levels, and 700m and 3300m medium- and long-sea outfalls. The sewer network has a further combined sewer overflow at 50.7647° N, 0.2747° E (WGS84) which discharges to the sea.

2.2.5 Flood risk

The study area consists of a complex catchment in which all forms of flood risk are present and interact in a manner that challenges conventional flood risk management approaches.

Surface water flooding occurs particularly in the lower parts of the catchment, where a combination of insufficient drainage system capacity and high groundwater impede conveyance of water away from properties. Surface water flooding also occurs in the north and west of the upper catchment. Furthermore, significant areas are vulnerable to groundwater flooding due to the underlying chalk geology.

The sewer system is also vulnerable to storm events and high groundwater, which can cause surcharging at low spots in the catchment. It is recognised that the sewer system has insufficient capacity for intense rainfall events and this has previously resulted in extensive surface water flooding in the town centre (Derek Finnie Associates, 2022). A contributing factor is the diversion of all flow from a historic, culverted watercourse (the Bourne Stream) into the surface water sewerage network, which results in the network capacity frequently being exceeded in storm events and leading to surface water flows.

Flood management challenges are exacerbated by tidal locking (as discussed in Section 2.2.3), as well as the tidal influence on groundwater levels (given the low-lying nature of the catchment). Shallow gradients also result in siltation of the system and affect the conveyance of flood water.

In total, Eastbourne has over 6000 properties at risk from surface water flooding under a 100-year rainfall event, 14,000 from coastal flooding, and 4000 from groundwater flooding. Fluvial flood risk is of a similar order, and the town is also at high risk of increased property sewer flooding due to planned residential growth (Derek Finnie Associates, 2022).



2.3 Sensor network and system monitoring

A dense network of sensors provides industry-leading spatio-temporal resolution data on water levels and rainfall depths across the catchment, as illustrated in Figure 3. Fluvial and borehole level gauges additionally provide water and air temperature data, although it should be noted that the air temperatures are captured by the level monitoring equipment predominantly for the purpose of auto-compensation for atmospheric pressure and temperature variation in the calculation of water levels from water pressure, and are not compliant with World Meteorological Organization standards and recommended practices for temperature monitoring.

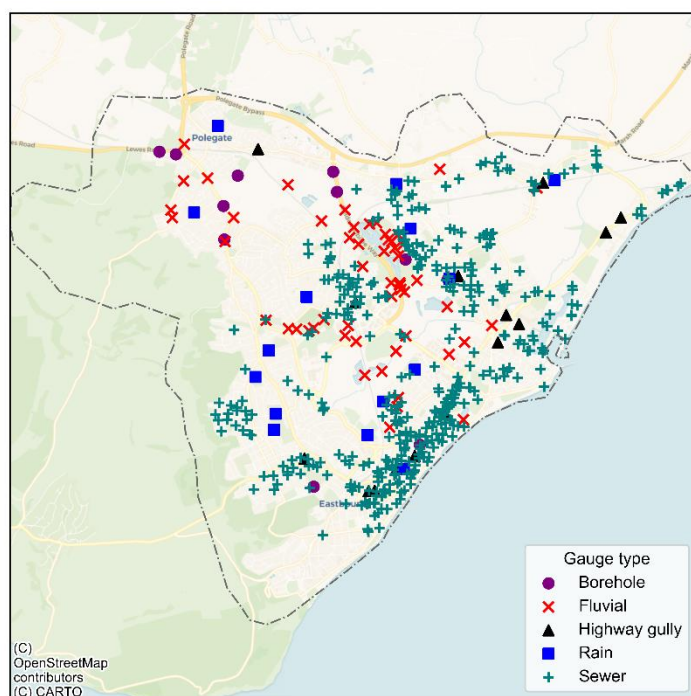


Figure 3: Distribution of Blue Heart- and Southern Water-owned gauges in the study area.

The transition to ‘smarter’ stormwater management provided by the Blue Heart project, including installation of sensors in a phased approach. The iterative steps were broadly in accordance with the framework presented by Sweetapple et al. (2023), and as such, availability of data across the catchment varies by date.

Phase 1 aimed to provide rapid installation of sensors at a small number of sites identified as key priorities in order that challenges could be identified and lessons learned before wider rollout in subsequent phases. It included monitoring of 7 boreholes, 7 fluvial levels and 5 surface water sewer levels, and was mainly completed in November 2021. Borehole and fluvial level monitoring was undertaken using In-Situ pressure transducers and telemetry (Rugged Troll 200 and VuLink) loggers with cellular connectivity, and for the purposes of testing and learning lessons (particularly with respect to battery



life), a range of different logging intervals (5-60 minutes) were initially implemented across different locations. The surface water sewer level monitoring was installed by Southern Water in February 2023 for the Blue Heart Project using radar level sensors (UDLive CatsEye Pixel II devices) with cellular connectivity and 15 minute logging intervals.

Phase 2 included installation of 51 water level sensors (KISTERS iLevel-GW pressure transducer and data logger) and 15 rain gauges (KISTERS TB7 tipping bucket) during 2024. These were sited to achieve a distribution across the catchment and capture locations of interest whilst also taking into account practical constraints such as accessibility, ease of maintenance, cellular coverage and likelihood of vandalism. Data from all Environment Agency gauges in the catchment (4 river level, 1 rain), were also integrated in this phase.

Under Phase 3, 15 highway gully monitors (UDLive CatsEye Pixel II devices) were installed in March 2025. Phase 3 also included integration of data from a further 393 sewer level monitors installed by Southern Water under their Network Digitisation 2022 Programme (Southern Water, 2024). These included MetaspHERE and UDLive CatsEye Pixel I and II devices, although the detail on the distribution of these is unavailable. A further 4 borehole level sensors (KISTERS iLevel-GW) were installed in January-February 2026.

A summary of the installed gauges is provided in Table 1.

Table 1: Summary of installed sensors in the Blue Heart case study catchment, including identification of measured parameters available for each.

Category	Ownership	Number of locations	Installation date range / date of first data	Device(s)	Device type	Logging interval (minutes)	Parameters				
							Water level or depth	Rainfall depth	Water temperature	Air temperature	Signal-to-noise ratio
Boreholes	Blue Heart	7	2022-07-13	In-Situ Rugged Troll 200 and VuLink	Pressure transducer	60	✓		✓	✓	
		4	2026-01-29 – 2026-02-02	KISTERS iLevel-GW	Pressure transducer	15	✓				
Fluvial system	Blue Heart	7	2021-11-09 – 2021-12-22	In-Situ Rugged Troll 200 and VuLink	Pressure transducer	5 – 60	✓		✓	✓	
		51	2023-12-06 – 2024-01-26	KISTERS iLevel-GW	Pressure transducer	5*	✓		✓	✓	



	Environment Agency	4	1999-07-13 – 2007-10-19	Unknown	Unknown	15	✓	
Sewer system	Southern Water	5	2023-02-23 – 2023-11-30	UDLive CatsEye Pixel II	Radar	15	✓	
		393	2022-06-28 – 2024-09-20	Mixed	Radar	15	✓	
Rain gauges	Blue Heart	15	2024-02-19 – 2024-07-31	KISTERS TB7 tipping bucket	Tipping bucket	15	✓	
	Environment Agency**	1	1991-01-12	Unknown	Unknown	15	✓	
Highway gullies	Blue Heart	15	2025-06-30	UDLive CatsEye Pixel II	Radar	15	✓	✓

* Air temperature is logged at a reduced interval of 30 minutes

** Data from Environment Agency gauges is already publicly accessible and is not reproduced in this publication (see Section 4 for further information)

2.4 Data collection, processing and storage

All sensors operate on independent logging systems with varying sampling intervals (typically 5-60 minutes, as outlined in Table 1). Raw data is communicated via cellular technology and stored by the telemetry providers; it is then accessed in .csv format via API.

Raw data files are structured with a ‘timestamp’ column, followed by a column for each parameter recorded at the corresponding gauge (as indicated in Table 1). Units are also indicated in the column headings. No conversions are applied to the raw data, and thus the units of water level and depth measurements are variable between gauge types – e.g. mm for highway gully water levels and metres or metres above ordnance datum (mAOD) for gauges in the fluvial network. Phase 1 borehole measurements are recorded as ‘depth to water’, i.e. illustrating an inverse relationship to water level. All units are identified in a gauge index file (index_gauge_data.csv), as well as in the corresponding data column heading (e.g. “Water Level [mAOD]”).

Availability of raw data across the monitoring period is illustrated in Figure 4. Data spans the period 9th Nov 2021 to 30th April 2026. There is a continuous period of 40 months with data available from at least 300 gauges (December 2022 onwards), and 26 months with data from at least 400 gauges (January 2024 to March 2026). The distribution of monitoring durations and raw data completeness across individual gauges is shown in Figure 5. Gauge-level monitoring periods range from 73 to 1581 days (median 1246 days), within which gauge downtimes range from 0.0 to 93.6 % (median 6.9%).

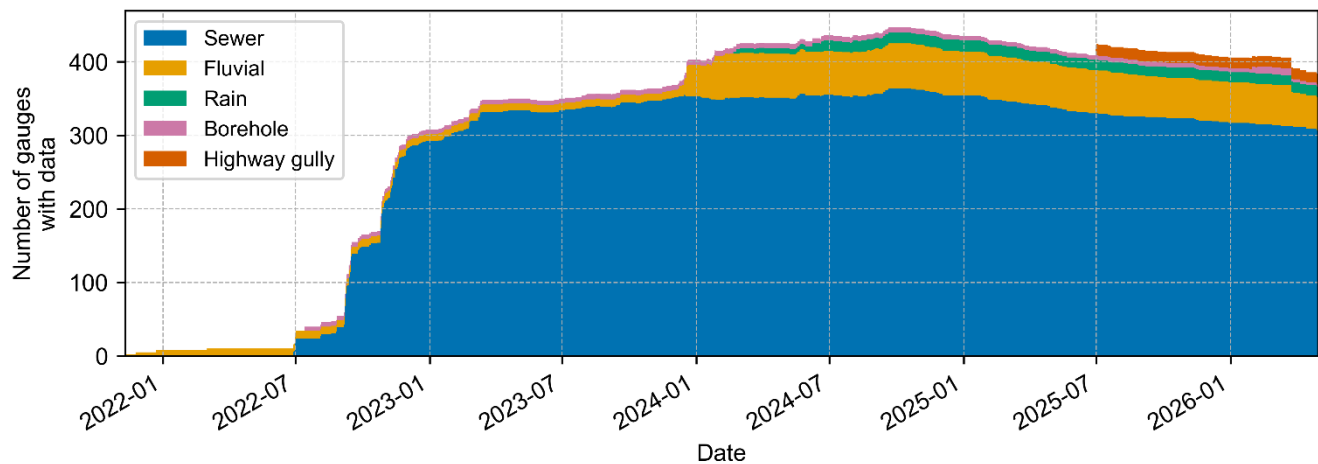


Figure 4: Number of gauges of each type with data available each day during the monitoring period (before quality checks and removal of any erroneous data)

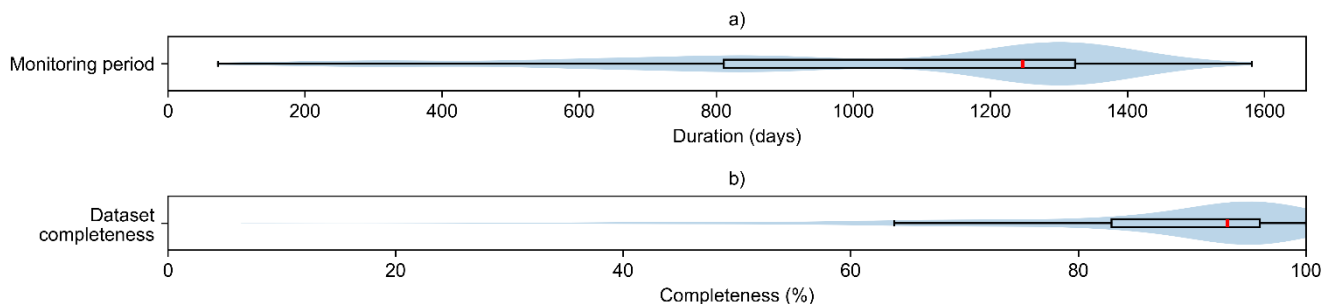


Figure 5: Distribution of raw data characteristics calculated for each gauge: a) Total duration of monitoring period (i.e. first to last measurement) and b) Dataset completeness (pre-quality control checks).

2.5 Data validation and quality control procedures

Identification of potentially erroneous data is vital, since random and systematic errors in the raw data may be present due to issues such as equipment or installation errors, discontinuities in power supply or data transmission, or human error (Clemens-Meyer et al., 2021). Preliminary, automated quality control checks have, therefore, been applied to all gauge data, and quality flags generated for each data point. To ensure that all raw data is retained in an unedited format, the quality flags are provided in separate files, but may be accessed in parallel to enable identification and removal of erroneous data before it is used in analyses.

Data quality checks are based on the following categories, with each data point classified according to the flags in Table 2:

- **Range:** Parameter- and gauge-specific limits are used to identify implausible or suspect values. This includes, for instance, flagging of negative rainfall or water depths; flagging radar-based measurements with a corresponding signal-to-noise ratio (SNR) below 10dB, and flagging air and water temperatures outside the plausible range for the



region. Additionally, the hydraulic for the catchment model includes topographical cross sections which provide estimated bed and bank levels for 38 of the 58 fluvial gauge locations. There is a degree of uncertainty in their accuracy (e.g. due to sediment deposition and erosion between the survey date and time of data collection); therefore, measurements are only flagged as ‘bad’ if they are more than 0.1m below the corresponding bed level, with those that are up to 0.1m below flagged as ‘suspect’. Further, given that no known major flooding occurred during the monitoring period, measurements that exceed the estimated bank level by more than 0.5m are also flagged as ‘suspect’.

- **Rate-of-change:** Hydrologically implausible or improbable step changes are identified using parameter- and gauge-specific thresholds; data corresponding to higher rates of change are flagged as ‘suspect’ and require further investigation. Different thresholds are applied to water levels recorded in different systems, due to their different dynamics: Groundwater levels, for instance, may only change slowly, and thus a threshold of 0.05m / hour is applied. Conversely, water levels in sewers and gullies may be subject to more rapid changes, and a threshold of 0.1m/minute is applied here. In the fluvial system, rates of change exceeding 1.0m / hour are considered highly improbable (whilst river level rises of several metres within hours are possible during severe events (Fileni et al., 2026), such extreme responses have not previously been observed in the case study catchment, and would warrant validation if found in the data). Temperature rarely changes extremely fast, and thus rates of change exceeding 3°C / hour and 5°C / hour for water and air temperatures respectively are considered suspect.
- **Persistence (flatline detection):** Timeseries are checked for unexpected, extended periods of stationary readings which, although theoretically possible, may indicate a sensor fault such as fouling. For water levels, persistent zeroes are plausible as the channel or sewer could be dry under some conditions, and thus only persistent non-zero values are flagged as ‘suspect’. Variable time thresholds for flagging of persistent values are applied, with 6 hours used for fluvial and sewer levels and temperatures, and 24 hours used for borehole water levels due to the slower dynamics of groundwater. Highway gully water levels are omitted from the persistence checks as, while static levels are commonly observed at all locations, the corresponding SNR values exhibit minor fluctuations and are high enough to indicate that the measurements are reliable. This may be attributed to sitting water below the outlet level.

Detailed criteria for each of the checks applied under these categories and the quality flag assigned to data that fails each check are provided in Table 4. In the case that multiple failure criteria are met for a data point, the ‘worst’ quality flag is used. Quality flags (Table 2) are provided in conjunction with labels (Table 3) that indicate the quality check failure(s) contributing to flag selection (adapted from Clemens-Meyer et al. (2021)): The code ‘3AD’ for instance, would indicate that a data point has been categorised as ‘bad data’ due to being above the feasible range and also exhibiting an excessively high rate of change.

No quality issues were flagged in sensor manufacturer-provided quality flags, and there were no known maintenance issues that would have resulted in untrustworthy data. More comprehensive data validation – for instance based on outlier detection, cross-sensor consistency checks (e.g. comparing upstream and downstream gauges), hydraulic and hydrological consistency



295 checks (e.g. based on water level responses to rainfall), sensor drift detection, and machine learning approaches for anomaly detection – remains an opportunity for future research.

Table 2: Quality flags used

Flag	Description
1	Quality control checks passed
2	Suspect data
3	Bad data

Table 3: Labels indicating reason for quality flag

Label	Description
A	Above feasible range
B	Below feasible range
C	Low SNR
D	High rate of change
E	Persistently constant value

300



Table 4: Quality control checks applied, including criteria for flagging and the flags used.

Category	Parameter	Criteria	Flag	Label
Range checks	Water level	Up to 0.1m below the estimated bed level*	2 – Suspect	B
		More than 0.5m above the estimated bank level*	2 – Suspect	A
		More than 0.1m below the estimated bed level*	3 – Bad	B
	Water depth	Corresponding SNR below 10 dB**	2 – Suspect	C
		Below 0mm	3 – Bad	B
	Sewer level	Below 0mm	3 – Suspect	B
	Depth to water	Below 0mm	3 – Bad	B
	Rainfall depth	Below 0mm	3 – Bad	B
		Above 50mm (equivalent to 200mm/hr intensity)	3 – Bad	A
	Water temperature	Below -2°C	3 – Bad	B
		Above 35°C	3 – Bad	A
	Air temperature	Below -20°C	3 – Bad	B
		Above 50°C	3 – Bad	A
Rate-of-change checks	Sewer level	Greater than 0.1m / minute	2 – Suspect	D
	Water depth and water level (fluvial)	Greater than 1.0m / hour	2 – Suspect	D
		Greater than 0.1m / minute	2 – Suspect	D
	Depth to water and water level (borehole)	Greater than 0.05m / hour	2 – Suspect	D
		Greater than 3°C / hour	2 – Suspect	D
	Water temperature	Greater than 5°C / hour	2 – Suspect	D
		Greater than 5°C / hour	2 – Suspect	D
Persistence checks	Sewer level	Identical, non-zero readings for more than 6 hours	2 – Suspect	E
	Water depth and water level (fluvial)	Identical, non-zero readings for more than 6 hours	2 – Suspect	E
		Identical, non-zero readings for more than 24 hours	2 – Suspect	E
	Depth to water and water level (borehole)	Identical readings for more than 6 hours	2 – Suspect	E
		Identical readings for more than 6 hours	2 – Suspect	E
	Air temperature	Identical readings for more than 6 hours	2 – Suspect	E

* Estimated bed and bank levels are available for 65% of fluvial gauges only

** SNR data is available for highway gully gauges only

Overall, 84.5% of data points pass all quality control checks (flag = 1), 15.1% are classified as ‘suspect’ (flag = 2), and 0.4% classified as ‘bad’ (flag = 3). This is skewed by a relatively small proportion of gauges containing a high proportion of flagged



305 data: At a gauge-level, the proportion of data with a high quality is typically much better (based on all data collected from each gauge individually, the median proportions of data flagged as ‘suspect’ and ‘bad’ are only 3.49% and 0.00% respectively). Figure 6 shows the gauge-level distribution of the percentage of data points assigned a quality flag of 1, separated by parameter. The top bar, for example, shows that the sewer level gauges have a median proportion of 94.5% of data with a quality flag of 1; however, there is significant variability, with one gauge having 0.0% of data with a quality flag of 1 (although this is
 310 predominantly due to ‘suspect’ levels below 0m and may still successfully indicate trends). Data quality for all other gauge types and parameters is more consistently high, with median values for the proportion of data with a quality flag of 1 of at least 97.8%.

Identification of data points as either ‘suspect’ or ‘bad’ is predominantly attributed to values being below the feasible range (82.3% of data points with a 2 or 3 flag) or failing persistence checks (16.7%). Only 2.0% are associated with high rates of
 315 change, 1.5% with above range values, and less than 0.1% with a low SNR.

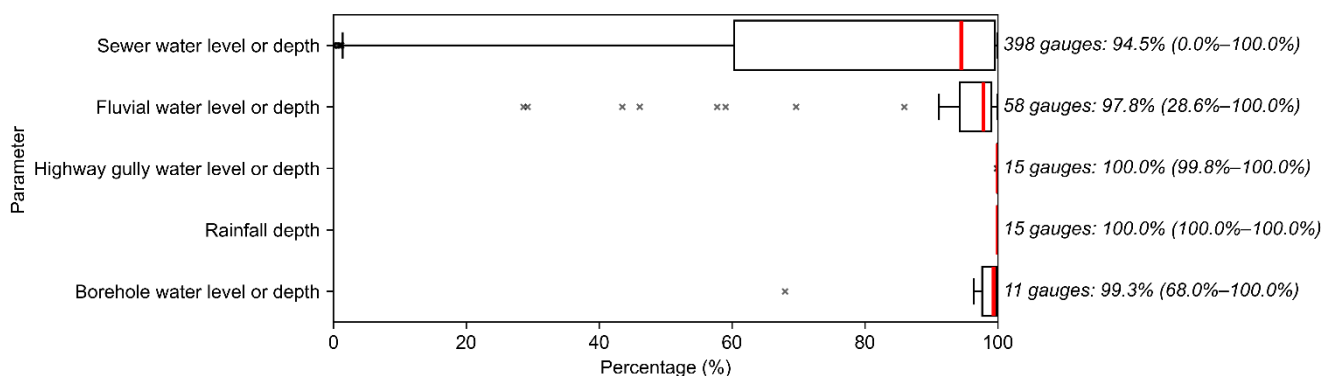


Figure 6: Distribution across all gauges of the gauge-level percentage of data with a quality flag of 1, separated by parameter. Labels to the right indicate the number of gauges at which the corresponding parameter is measured, and the median, minimum and maximum percentage of the data at each gauge assigned a quality flag of 1.

320 3 Results and discussion

3.1 Data summary

This section provides a high-level illustration of the data collected, and includes only data that passed the set of quality control checks.

Daily total rainfalls recorded at each gauge are shown in Figure 7; days with less than 95% data completeness are omitted, and
 325 data gaps are identified. Peak 15 minute rainfall intensities in excess of 50 mm/hr were recorded at 5 locations (33%), although these were short duration and 24 hour totals exceeding 50 mm were only captured at 4 gauges. Across the catchment, a maximum 15 minute intensity of 75mm/hr captured at one gauge.

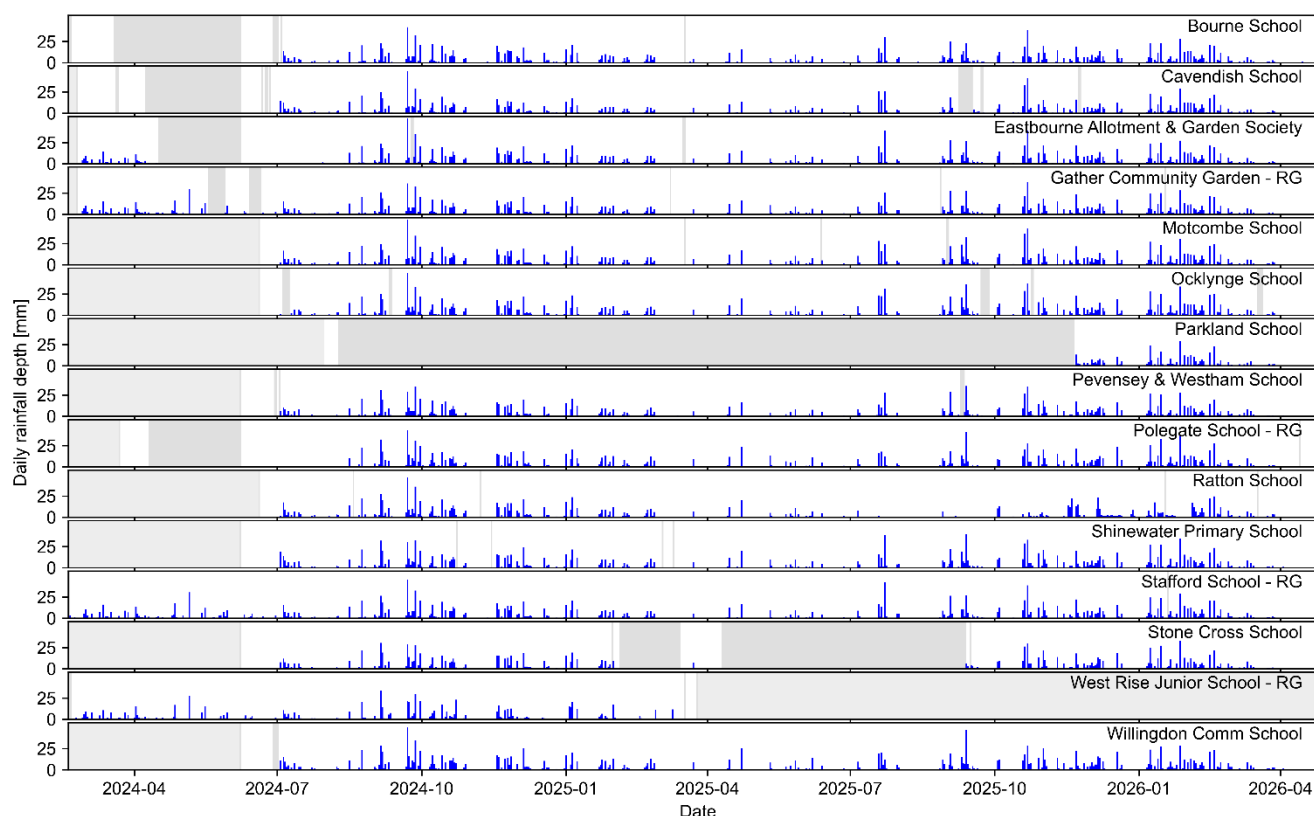


Figure 7: Total daily rainfall recorded at each rain gauge. Grey shading indicates data gaps.

330 Example water level timeseries data collected from fluvial gauges, boreholes, highway gullies and sewers are shown in Figure 8, Figure 9, Figure 10 and Figure 11 respectively. For fluvial gauges, highway gullies and sewers, only small subsets of the gauges are shown to avoid excessive overlaps and obscuration (chosen according to the alphabetical ordering of gauge names). Given significant variability in the baseline water levels for boreholes and highway gullies, levels for these are plotted as deviation from the gauge-specific minima, rather than absolute levels.

335 Figure 8 suggests broad alignment in the onset of significant water level disturbances between fluvial gauges, but with variability in the extent to which the water level rises and rate at which levels recede. The range of water levels observed at a single gauge also varies significantly between locations (minimum 0.27m, median 0.96m, maximum 2.87m).

The borehole data (Figure 9) shows that there is significant variability in groundwater levels across the catchment. Although absolute levels are not comparable (levels shown are relative to the corresponding borehole's minimum), it is evident that

340 some locations are subject to significantly greater variability. Based on the gauges with at least one year of data, for example, minimum range of 0.98 m was observed, and a maximum of 4.48 m.

The example highway gully data in Figure 10 shows notable differences between gullies, particularly in the baseline water level (across all gauges, the minimum value ranges from 1 mm to 747 mm). There are frequent, short-duration spikes, assumed



to be associated with the rainfall runoff response, which appear to be more frequent during winter months (mid October to mid March). This is consistent with typical seasonal increase in rainfall. The range of typical levels observed at each gauge also appears to vary significantly between locations (1st-99th percentile ranges vary from 31 mm to 535 mm). Potential anomalies, where water level drops below the typical baseline for a location are also observed; while such drops are plausible (e.g. due to gully cleaning), they require further exploration.

For sewers, the example gauges plotted in Figure 11 indicate that levels can fluctuate rapidly throughout the year, but typically remain within gauge-specific upper and lower bounds (likely determined by the pipe and/or chamber characteristics). The levels here appear to be less rainfall driven than at the fluvial and highway gully gauges, given the continuous fluctuations at some gauges, suggesting that the ‘foul’ element of the sewer flow dominates; however, further research is needed to characterise individual gauges and better understand their different rainfall responses.

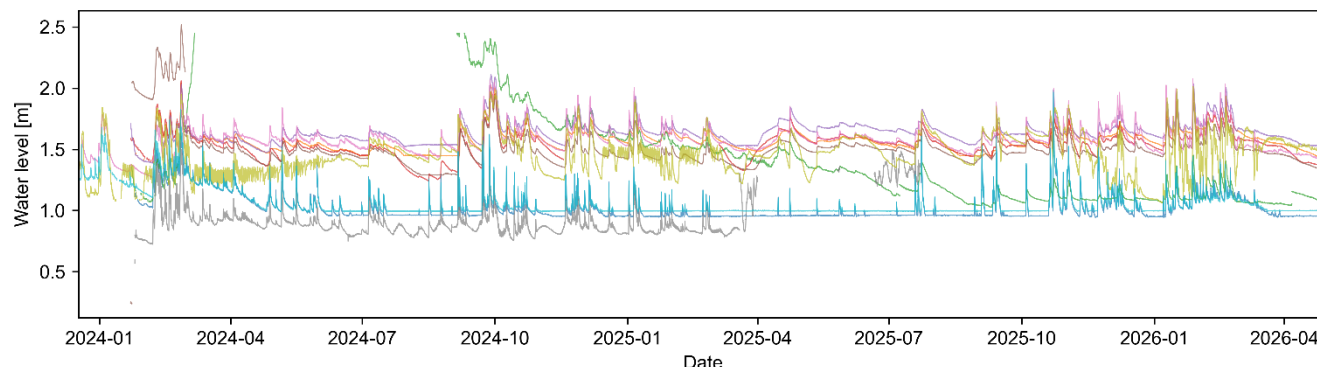


Figure 8: Water level data collected from 10 example fluvial gauges.

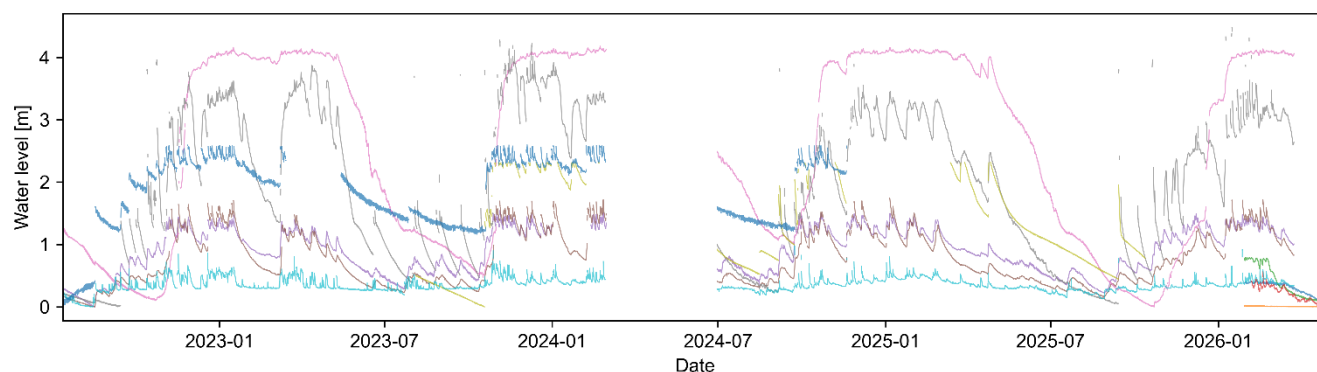


Figure 9: Water level data collected from boreholes, reported relative to minimum level observed at each gauge.

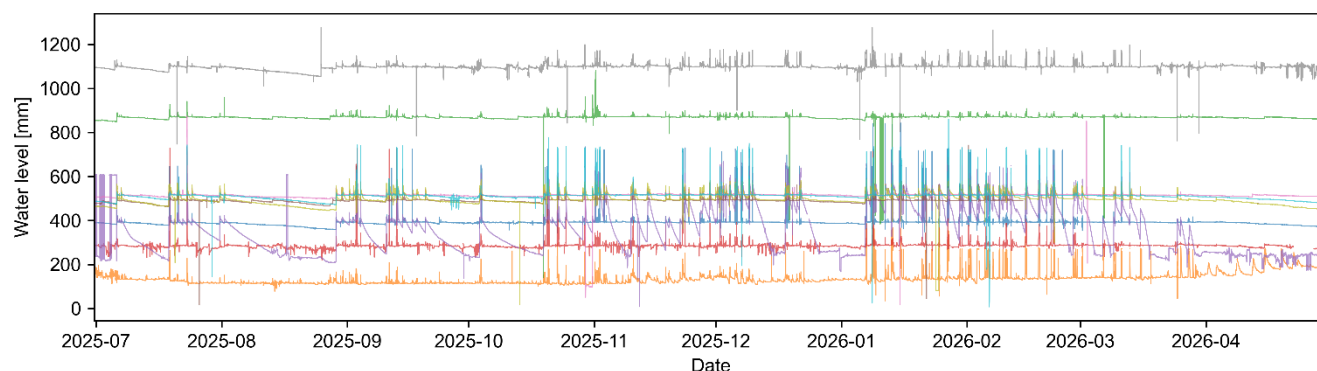


Figure 10: Water level data collected from 10 example highway gullies, showing deviation from the minimum recorded level at each gauge.

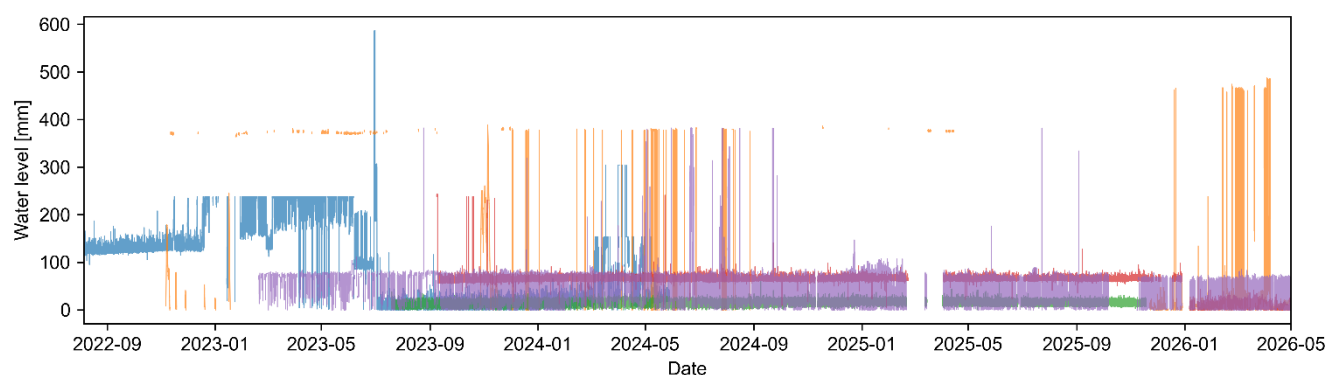


Figure 11: Water level data collected from 5 example sewer gauges.

Example air and water temperature timeseries data are shown in Figure 12 and Figure 13 respectively. Both exhibit clear seasonal trends, with water temperatures fluctuating within a narrower range both seasonally and daily than air (as expected).

Across all gauges with at least a year of data, air temperatures exhibit a median 1st-99th percentile range of 28.3 °C; Water temperatures from fluvial gauges exhibit a median 1st-99th percentile range of 15.8 °C, and for borehole gauges this reduces significantly to 2.2 °C.

There appears to be a high degree of consistency between air temperature patterns in measured in different locations across the catchment, with an interclass correlation coefficient of 0.92 and a mean inter-gauge standard deviation of 1.89 °C (maximum 6.02 °C). Fluvial water temperatures are less consistent, with an interclass correlation coefficient of 0.77.

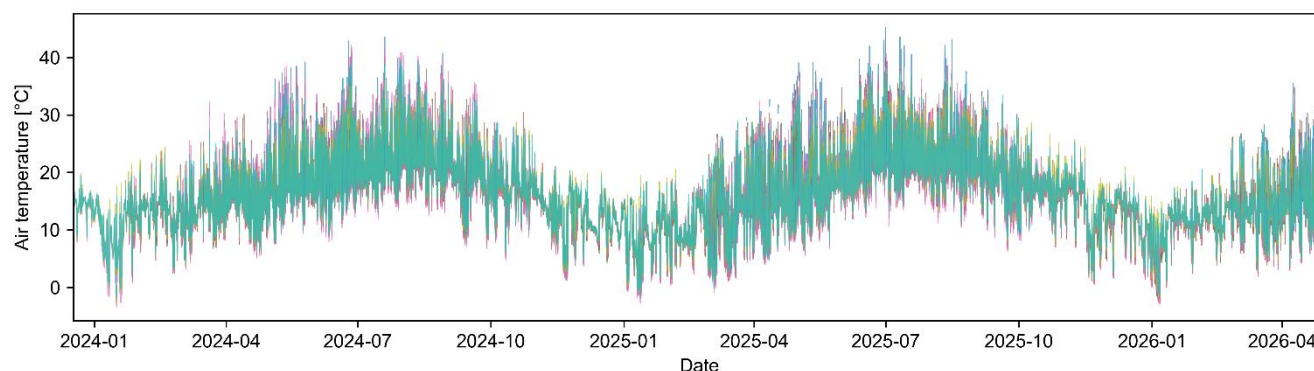


Figure 12: Air temperature data collected from 10 example gauges.

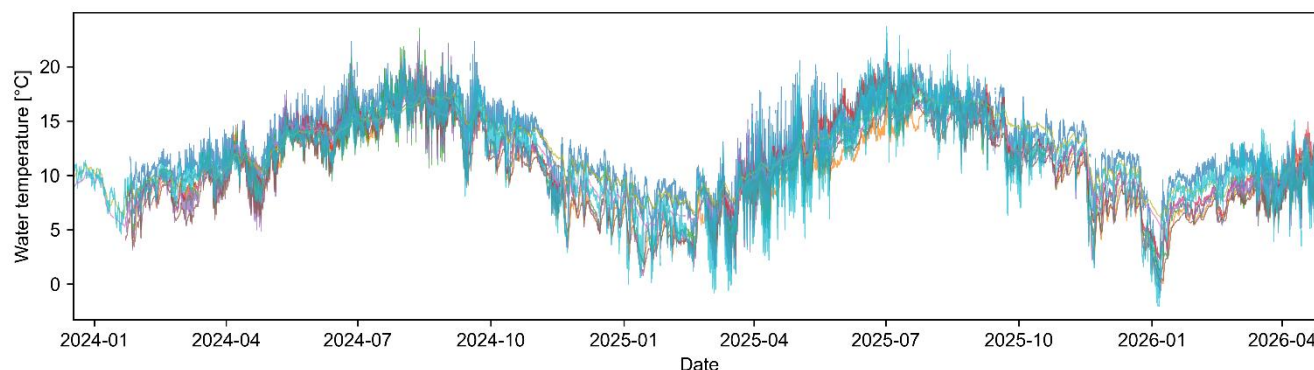


Figure 13: Water temperature data collected from 10 example gauges.

375 3.2 Research opportunities and potential applications

Long-term, high-resolution monitoring has been identified as critical for advancing understanding, testing new methodologies and improving urban water management in complex catchments (e.g. Blumensaat et al., 2025; Hoard et al., 2020; Oh and Bartos, 2025; Roosipuu et al., 2023), yet integrated, high-resolution monitoring remains rare in urban hydrology and urban drainage research. This dataset offers an unusually high-density, multi-component view of an urban catchment that spans the sewer network, fluvial channels, groundwater, highway gullies and rainfall, with additional supporting temperature observations, and offers the potential for a range of applications. Of particular interest is aiding the transition to ‘smarter’ stormwater management, which has been recognised as a promising approach to tackling the challenges arising from stressors such as climate change and urbanisation (Sweetapple et al., 2023; Webber et al., 2022). At present, real-world development and implementation of such data-driven, ‘smart’ stormwater management systems is hindered by a variety of technical and socio-economic barriers and challenges, ranging from decision making (e.g. selecting what and how much data to collect, development of control algorithms, and performance evaluation methods) and addressing uncertainties (e.g. uncertainty in sensor measurements, performance uncertainty and knowledge gaps), to addressing social barriers such as lack of confidence,



resistance to change, insufficient knowledge and governance (Sweetapple et al., 2023). However, this dataset provides a unique opportunity to investigate urban hydrological processes across spatial and temporal scales that are rarely captured within a single monitoring programme, and is well-positioned to help address many of the previously identified challenges. Furthermore, Open benchmark datasets are essential for accelerating methodological innovation, improving transparency and enabling reproducible research, yet, large, openly available urban hydrological datasets remain comparatively rare, particularly those containing extensive sewer system observations. The dataset, therefore, has substantial value as a benchmark resource for method development and comparison studies. By making these observations publicly available, the dataset can enable collaboration between different disciplines and areas of expertise – for instance, hydrologists, engineers, computer scientists, data scientists, infrastructure operators and flood risk practitioners – with the potential to enhance understanding of complex urban water systems and support the development of more resilient built environments.

A number of research opportunities enabled by this dataset and potential applications are discussed in further detail in the remainder of this section, although this is not an exhaustive list.

3.2.1 Machine learning, artificial intelligence and digital twin development

Machine learning, AI and digital twin approaches have the potential to be transformative in urban stormwater management, with possible applications including predictive modelling, design and simulation, infrastructure optimisation and decision support (Rahman et al., 2026). Whilst these approaches are being increasingly studied (e.g. Ramovha et al., 2024; Santos et al., 2025; Sharifi et al., 2024), many methods are currently developed and tested using simulated datasets (e.g. El Ghazouli et al., 2022; Zheng et al., 2026) and an absence of comprehensive datasets remains a fundamental barrier to the widespread adoption of AI for stormwater management (Rahman et al., 2026).

Furthermore, the current (lack of) availability and reliability of existing models for real-time control that integrate with real-time data is recognised as a challenge (Xu et al., 2021), and datasets that facilitate further development in this area may be critical to the transition to ‘smarter’ stormwater management systems and realisation of the benefits they offer. It has also been suggested that models used for real-time control need not necessarily be as complex as those currently used for simulation (Mullapudi et al., 2017), thus indicating an opportunity to build simpler, data-driven models that interface with real-time data. The data presented in this study could be used to help address the current scarcity of suitably detailed real-world observational data, with the extensive scale of the dataset providing opportunities for development and benchmarking of machine learning and AI models for applications such as hydraulic state estimation, anomaly detection, sensor fault identification and flood prediction. The dataset is particularly valuable because it contains observations from interconnected components of the urban water system, as well as geospatial information on the network layout and connectivity, enabling exploration of complex system dynamics via machine learning and AI approaches without the need to explicitly model processes such as groundwater interactions, infiltration or operational control.

The multi-component nature of the dataset (including rivers, sewers, groundwater, highway gullies and rainfall) also provides interesting foundation for research into digital twins for urban flood management, infrastructure resilience and adaptive asset



operation, given the digital twin requirements for continuous observations to update model states and evaluate system performance in near real time.

3.2.2 Understanding spatio-temporal variability in urban hydrological processes and responses

Urban hydrological processes can exhibit high variability across small spatial and temporal scales, with rainfall, runoff generation, sewer response, and flood propagation potentially differing substantially between neighbouring sub-catchments due to variations in topography, land use, drainage connectivity, and infrastructure characteristics (Cristiano et al., 2017; Macdonald et al., 2022). Figure 7, for instance, shows rainfall captured at an example rain gauge in the case study catchment during an example storm event, along with the water level response at 19 fluvial, 45 sewer, 1 highway gully and 2 borehole (groundwater) monitoring locations within a 1.25 km radius, illustrating the extent to which the rainfall responses varies within urban water features across a small area. This evidence highlights the importance of sensor-typology and sensor-placement in the design of data-driven evaluation approaches, and aligns with the theory that causal relationships that appear present between two specific sensors might not be scalable across the wider catchment.

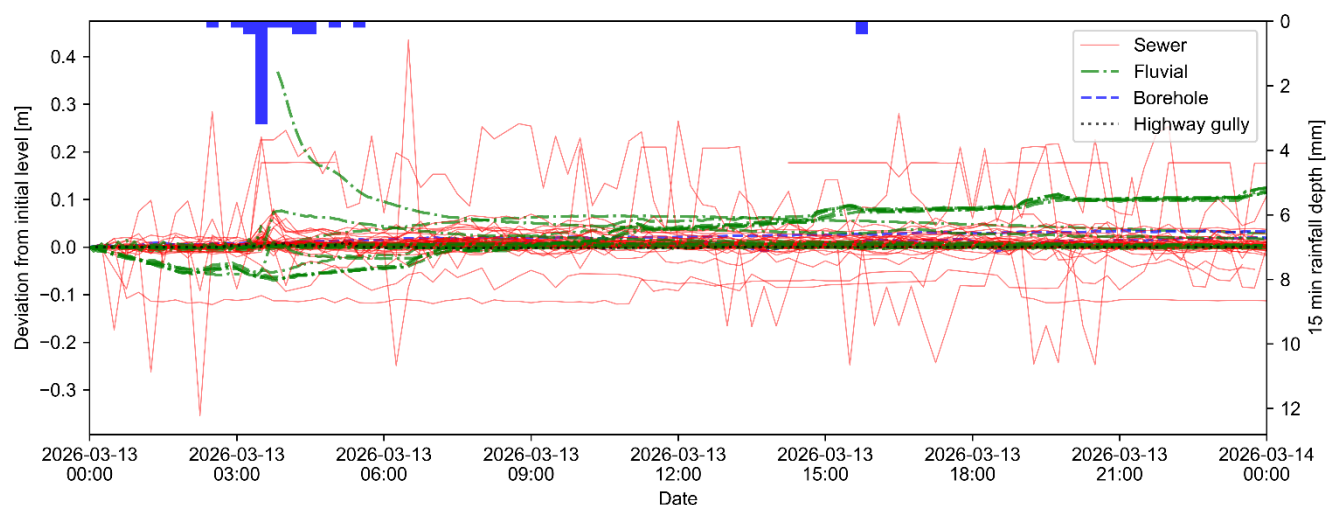


Figure 14: Example rainfall and water level data collected at gauges within a 1.25 km radius of a rain gauge during and following a storm event. Changes in water level shown are relative to the level at the corresponding gauge at the start of the event.

The value of high-resolution monitoring for identifying urban hydrological variability that may remain undetected in conventional, sparse monitoring networks has been demonstrated previously (e.g. Blumensaat et al., 2025; Hoard et al., 2020); however, previous studies have been limited (in comparison with this study) in their scale and/or the system components covered. By providing high spatial and temporal resolution observation data encompassing rainfall depths and water levels in the sewer network, fluvial network, highway gullies and boreholes, this study offers the opportunity to explore such heterogeneities further and to improve understanding of the interactions between rainfall, groundwater processes, river hydraulics and sewer network performance at a high resolution. Such analyses may also enhance understanding of spatial



variability in attenuation, response time, sub-system connectivity, and thresholds for flooding in different sub-systems. Improved understanding of these processes is essential for assessing urban resilience to increasingly frequent extreme rainfall events, and may also support the development of more representative model structures and parameterisations. Better understanding of spatio-temporal variability in rainfall and the water system response could also inform more efficient sensor network design methodologies (discussed further in Section 3.3.2).

3.2.3 Improving sensor network design and addressing infrastructure limitations

When developing catchment monitoring systems, it is necessary to decide what type of data should be collected, where and how frequently. Robust sensor network designs typically aim to reduce uncertainty in predictions at ungauged locations (although socio-economic and operational objectives may also be pertinent), and are essential during crisis events to inform real-time decision making, identify flood risks and enable rapid and targeted responses (Oh and Bartos, 2025). As illustrated by the variability in rainfall responses at gauges within a 1.25 km radius shown in Figure 7, however, a single gauge cannot necessarily be assumed representative of conditions in even relatively close locations.

It is acknowledged that the amount of data required may vary between applications, with more data not necessarily being better (for instance, due to governance, cost and resource constraints (Oh and Bartos, 2025; Sweetapple et al., 2023)); however, determining the optimal can be challenging since the relationship between resolution and benefits will be specific to the application (Eggimann et al., 2017) – for example, a sensor network optimised for flood warning may differ from one designed for control applications. In the context of this project, a key goal was to obtain maximum data from a fixed budget during a long term monitoring study, for the purpose of better understanding the system. However, in future settings, optimal sensor placement is a crucial challenge to address if sustainable hydrological monitoring networks are to be achieved at scale (Oh and Bartos, 2025). Furthermore, the design and cost effective implementation of sensor networks has been identified as one of the potential issues hindering the transition to ‘smarter’ stormwater management systems (Sweetapple et al., 2023).

Because the sensor network upon which this paper is based is both dense and heterogeneous, it could be used in further research to retrospectively test strategies for optimal sensor placement, redundancy and network thinning – for example, exploring which sensors contribute most to reducing model or forecast uncertainty in different sub-systems, and building on recent work on data-driven design of hydrological monitoring networks (e.g. Oh and Bartos, 2025). Likewise, researchers might explore the integration of further data such as satellite monitoring or telemetry gathered from mobile devices or autonomous vehicles to set out alternative sensor-placement strategies.

Work on optimal sensor monitoring design could also explore methods for addressing infrastructure limitations that have been identified as further barriers to the adoption of ‘smarter’ stormwater systems, such communication range and technological constraints (Sweetapple et al., 2023) – for instance, incorporating restrictions on sensor placement to emulate communication range limitations. Additionally, approaches to minimising the amount of data transmitted and/or the frequency of transmission (for example, to save battery life in off-grid locations or address bandwidth limitations) whilst still providing sufficient data



475 that its end applications are not adversely affected could be explored – this might include, for example, investigating the impact of different strategies for dynamic sampling, considering variability between different types of location or water sources.

3.2.4 Characterising groundwater-sewer interactions and sewer system performance

This dataset has the potential to provide detailed understanding of interactions between different hydrological components and responses in the urban water system – for example, interactions between groundwater and sewer levels. At present, sewer
 480 inflow and infiltration represent major operational challenges for water utilities because they can increase hydraulic loading, reduce treatment efficiency and elevate flood risk and pollution risks (Almeida et al., 2025; Wang et al., 2024). However, detecting and quantifying inflow and infiltration remains difficult due to the spatial extent of urban drainage networks and limited availability of observational data (Zhang et al., 2025).

The extensive sewer monitoring data provided in this publication for 398 locations, along with the associated rainfall and
 485 borehole monitoring presents opportunities to investigate temporal and spatial patterns in the sewer network response to rainfall and fluctuations in groundwater level, including identification of surface inflow pathways and hydraulic bottlenecks. It could also enable investigation into bidirectional interactions between groundwater and buried drainage infrastructure, including infiltration into sewers during high groundwater conditions and exfiltration where pipes are surcharged, and development and testing of associated, transferable assessment methodologies. Understanding these exchanges is important for operational
 490 efficiency, flood risk and water quality management (e.g. Ellis and Revitt, 2002; Karpf et al., 2011).

Recent studies have also highlighted the need the real-world datasets to test emerging approaches to sewer performance monitoring, anomaly detection and inflow and infiltration identification (e.g. Ge et al., 2024). Building upon the sensor network design opportunities discussed in Section 3.3.2, the availability of observations from 398 sewer monitoring locations in this dataset creates opportunities to evaluate how monitoring density influences sewer system understanding and operational
 495 decision making.

3.2.5 Data-driven performance assessment, asset management and decision making

At present, there are recognised uncertainties and knowledge gaps around stormwater system performance. This includes, for instance, understanding of how multiple real-time controlled assets interact at a catchment level (Schmitt et al., 2020) and potential dichotomy between local best management practices and larger scale impacts (Ibrahim, 2020; Mullapudi et al., 2018),
 500 as well as the more fundamental question of what performance evaluation methods are appropriate (considering, for example, what performance metrics should be used and what evaluation period) (Sweetapple et al., 2023). With urban flood risk expected to increase under future conditions, there is growing necessity for evidence-based approaches to asset management and decision making to provide adaptation and resilience and data from this case study could support assessment of infrastructure under a range of meteorological and hydrological conditions, including exploration of performance at different scales and
 505 based on different metrics.



The dataset may also be used to identify critical assets and system vulnerabilities - for instance, water level data from sewer and highway gully sensors could reveal the signatures of issues such as partial or full blockage, sediment accumulation or structural defects. More importantly, it could be used in the systematic development and validation of transferable, data-driven indicators for asset condition and methodologies for decision making and maintenance prioritisation.

510 3.2.6 Design and evaluation of smart stormwater system control strategies

Development and performance of control algorithms for stormwater management remains a challenge (Sweetapple et al., 2023), despite a growing body of research focused on stormwater system real-time control. At present, performance optimisation is typically achieved through control of a small number of system components at a local scale (Kändler et al., 2020), but the effectiveness of such simple strategies when applied to distributed stormwater systems is unclear. Developing
515 rules for coordinated and effective system-level control is more difficult, however (Bowes et al., 2021), and there remains a need for further research into the control of multiple-components at a catchment level (Webber et al., 2022).

Asset sensing, data collection and communication, and data analytics and integration are all essential prerequisites for online control and optimisation, either at an asset or a network level (Webber et al., 2022). It is proposed that the continuous rainfall, fluvial level, sewer level, borehole and highway gully data provided in this dataset could be used to develop and test control
520 strategies for storage operation, throttling and diversion within the stormwater system. Previous studies have shown that dense sensor networks can support real-time monitoring and control of urban drainage, providing measurable reductions in flood risk (Roosipuu et al., 2023), and this dataset provides the hydrometric foundation needed to explore control options under real conditions.

3.2.7 Development of real-time urban flood warning systems and prediction tools

Recent research on sensor-based urban flood monitoring demonstrates the potential for high resolution detection of flood onset
525 and propagation in critical locations (Arshad et al., 2019; Jang and Jung, 2023; Najafi et al., 2024). The high spatial and temporal resolution data provided in this publication could enable existing methods to be tested and generalised for a United Kingdom catchment, as well as facilitating the exploration of new approaches that utilise data from different sub-systems (including rain gauges, the fluvial network, boreholes and highway gullies) to generate highly-localised flood warning
530 thresholds, nowcasting tools and anomaly detection algorithms. This might include, for example, exploring whether current radar forecast-based approaches for downstream flood risk projection can be improved upon with the use of real-time rain gauge data.

3.2.8 Ecohydrological and thermal regime studies

Although previous research has explored connections between rainfall and water temperature (e.g. Li et al., 2025a; e.g. Subehi
535 et al., 2010), few studies have been identified that combine in-situ urban water temperature, rainfall and water level data. The associated air and water temperature data included in this dataset, therefore, opens the opportunity for novel ecohydrological



and thermal regime studies – for example, joint analysis of rainfall, water level and temperature data could improve knowledge of how hydrological events, stormwater inflows and air temperature extremes affect the river temperature at short timescales and support the development of coupled hydrological-thermal models for informing management of urban water quality and ecological health. Such analyses could improve understanding of ecological stresses and the potential impact of heatwaves in urban waterbodies.

3.2.9 Addressing social barriers to stormwater management change

A range of socio-economic barriers are currently hindering change in stormwater management practices, including unfamiliarity and lack of trust in novel technologies (Frantzeskaki, 2019), risk aversion (Eggimann et al., 2017), resistance to change (Roy et al., 2008) and a lack of guidance (Naughton et al., 2021). There is potential for this dataset to provide improved trust in the capability of sensor and telemetry technologies, based on evidence provided by (for instance) analysis of the reliability of the sensors over an extended period. Any further benefits of this dataset realised by other research applications (as discussed in the remainder of this section) could also help to improve confidence in the use of ‘smart’ technologies in stormwater management, and contribute to the development of improved guidance for implementation. Linked to this, the premises of ‘trust’ and ‘familiarity’ are connected, in that humans are biased towards familiar options: To help build trust, the presence and availability of data could continue to be used to increase awareness within the catchment of water and climate adaptation challenges. Furthermore, social science-led studies that make use of data, for example by testing how local residents respond to information when provided in different formats can also be unlocked.

4 Code and data availability

Data are available from Zenodo at <https://doi.org/10.5281/zenodo.20699634> (Sweetapple et al., 2026a). The repository contains the following:

- A ‘*README.md*’ file, containing information on the repository contents and structure
- Separate folders for the following categories of data:
 - **Catchment shape files** (‘*1_catchment_shape_files/*’): Shape files providing the catchment outline; elevation data and infiltration data.
 - **Network shape files** (‘*2_network_shape_files/*’): Shape files providing attributes of ‘link’ components in the sewer and fluvial systems (including conduits, river reaches, weirs, sluices, valves and screens), and geometry and attributes of ‘node’ components in the sewer system (including manholes, outfalls and storage). Metadata for all links and nodes is also provided in .csv format.
 - **Borehole data** (‘*3_borehole_telemetry_data/*’): .csv files with raw timeseries data from boreholes (one file per gauge), plus quality codes (one file per gauge).



- **Fluvial data** ('4_fluvial_telemetry_data/'): .csv files with raw timeseries data from the fluvial system (one file per gauge), plus quality codes (one file per gauge).
- **Sewer data** ('5_sewer_telemetry_data/'): .csv files with raw timeseries data from the sewer system (one file per gauge)
- **Highway gully data** ('6_highway_gully_telemetry_data/'): .csv files with raw timeseries data from the highway gullies (one file per gauge), plus quality codes (one file per gauge).
- **Raingauge data** ('7_raingauge_telemetry_data/'): .csv files with raw timeseries data from the rain gauges (one file per gauge), plus quality codes (one file per gauge).
- An index of all gauge locations ('index_telemetry_data.csv'), including the following information for each location:
 - Name
 - Latitude and longitude
 - Gauge type
 - Paths for corresponding raw data and quality code files
 - List of parameters measured and list of corresponding parameter units
 - Owner of the gauge
 - Level above which flooding occurs (highway gullies only)
 - Bed and bank level estimates (select fluvial gauges only)

Example scripts for accessing, viewing, processing and analysing the data are available from Zenodo at <https://doi.org/10.5281/zenodo.21390823> (Sweetapple et al., 2026b). This includes scripts for the validation and quality control procedures presented as well as for replication of all plots and statistics presented in this paper.

Additional (already publicly accessible) data provided by the Environment Agency is not duplicated in this dataset publication, and can instead be obtained directly from the Hydrology Data Explorer (<https://environment.data.gov.uk/hydrology>) (DEFRA, 2024). Details of relevant gauges within the study area are provided in Table 5, below.



590 **Table 5: Summary of Environment Agency operated gauges in the Blue Heart case study catchment with data openly available from the Hydrology Data Explorer (DEFRA, 2024).**

Category	Name	Station reference	Date of first data	Logging interval (minutes)	Parameters	
					Water level or depth	Rainfall depth
Fluvial system	Langney	E7280	1999-07-13	15	✓	
	Wannock Mill	E7040	2007-07-01	15	✓	
	Fence Bridge Gate	E23500	2007-10-19	15	✓	
	Pevensey Bridge	E7321	2007-07-01	15	✓	
Rain gauges	Pevensey	E7295	1991-01-12	15		✓

5 Conclusions

This publication presents a large scale, high resolution observational dataset centred on Eastbourne (south east England) which was collected under the guidance of University of Exeter’s Centre for Water Systems in the Blue Heart project. The data integrates rainfall, fluvial water levels, groundwater dynamics, highway gully responses and sewer levels across an urban catchment. The monitoring network comprises of 497 sensors with an average density exceeding six sensors per square kilometre, which represents a substantial increase in spatial resolution relative to existing national hydrometric networks. The dataset includes over 66.6 million data points, and provides an unprecedented empirical basis for understanding interactions between fluvial, pluvial, groundwater and sewer processes. The dataset is further enriched with metadata describing the sensor network, shape files providing detail on the catchment topography and layout of the fluvial and sewer systems, and additional air and water temperature data providing contextual environmental conditions.

Long-term, high resolution monitoring is critical for advancing understanding, and by making these data openly available along with example processing and analytical scripts, this work aims to lower barriers to entry for researchers and practitioners seeking to analyse complex urban hydrological systems. The dataset provides a valuable benchmark for exploring the challenges associated with transitioning towards smarter, sensor-enabled and data-driven stormwater management, and offers a unique opportunity to investigate urban hydrological processes across spatial and temporal scales that are rarely captured within a single monitoring programme. A wide range of specific applications research challenges that may be addressed using this data have been discussed, with key examples including:

- Development and benchmarking of machine learning, AI and digital twin models for applications such as flood prediction, hydraulic state estimation, anomaly detection, sensor fault identification, and adaptive asset operation.



- Better understanding of interactions between rainfall, groundwater processes, river hydraulics and sewer network performance, including spatial-temporal variability in responses.
- Improving methodologies for sensor network design, considering optimal sensor placement and temporal data collection strategies for different applications.
- Testing methodologies for sewer system performance monitoring and anomaly detection, and providing a better understanding of interactions between sewer levels, groundwater levels and rainfall.
- Supporting systematic development of transferable, data-driven, approaches for performance assessment, asset management and decision making, including exploration of performance at different scales, identification of system vulnerabilities and prioritisation of maintenance.

620 However, it is anticipated that uses will extend beyond those that have been identified here.

6 Author contributions

All authors conceptualised the publication. CS curated the data and prepared the example scripts. CS prepared the original draft manuscript; PM and TI reviewed and edited the manuscript; and CS revised the paper and prepared the final submission, which was approved by all authors. TI and PM supervised the project.

625 7 Acknowledgements

This research was completed for the ‘Blue Heart’ project and funded by Defra as part of the £200 million Flood and Coastal Innovation Programmes which is managed by the Environment Agency. The programmes will drive innovation in flood and coastal resilience and adaptation to a changing climate.

630 The authors gratefully acknowledge Southern Water for providing access to their sewer monitoring data and for granting permission to share the dataset publicly.

8 Competing interests

The authors declare no competing interests.

9 References

- Almeida, M. do C., Brito, R. S., and Jorge, C.: Inflows into Wastewater and Stormwater Systems: Sources, Causes, and Assessment, *Water (Basel)*, 17, 1082, <https://doi.org/10.3390/w17071082>, 2025.
- Arshad, B., Ogie, R., Barthelemy, J., Pradhan, B., Verstaavel, N., and Perez, P.: Computer Vision and IoT-Based Sensors in Flood Monitoring and Mapping: A Systematic Review, *Sensors*, 19, 5012, <https://doi.org/10.3390/s19225012>, 2019.



- Berndtsson, J.: Green roof performance towards management of runoff water quantity and quality: A review, *Ecol. Eng.*, 36, 351–360, <https://doi.org/10.1016/j.ecoleng.2009.12.014>, 2010.
- 640 BGS: Geology Viewer, <https://geologyviewer.bgs.ac.uk/>, 2026.
- Blue Heart: Flood awareness and resilience in Eastbourne and Southern Wealden, <https://www.blueheart.org.uk/>, 2025.
- Blumensaat, F., Bloem, S., Ebi, C., Disch, A., Förster, C., Maurer, M., Rodriguez, M., and Rieckermann, J.: The UWO dataset – long-term observations from a full-scale field laboratory to better understand urban hydrology at small spatio-temporal scales, <https://doi.org/10.5194/essd-2024-47>, 2 January 2025.
- 645 Bowes, B. D., Tavakoli, A., Wang, C., Heydarian, A., Behl, M., Beling, P. A., and Goodall, J. L.: Flood mitigation in coastal urban catchments using real-time stormwater infrastructure control and reinforcement learning, *Journal of Hydroinformatics*, 23, 529–547, <https://doi.org/10.2166/hydro.2020.080>, 2021.
- Cea, L., Sañudo, E., Montalvo, C., Farfán, J., Puertas, J., and Tamagnone, P.: Recent advances and future challenges in urban pluvial flood modelling, *Urban Water J.*, 22, 149–173, <https://doi.org/10.1080/1573062X.2024.2446528>, 2025.
- 650 Clemens-Meyer, F. H. L. R., Lepot, M., Blumensaat, F., Leutnant, D., and Gruber, G.: Data validation and data quality assessment, in: *Metrology in Urban Drainage and Stormwater Management: Plug and Pray*, IWA Publishing, 327–390, https://doi.org/10.2166/9781789060119_0327, 2021.
- Cristiano, E., ten Veldhuis, M.-C., and van de Giesen, N.: Spatial and temporal variability of rainfall and their effects on hydrological response in urban areas – a review, *Hydrol. Earth Syst. Sci.*, 21, 3859–3878, [https://doi.org/10.5194/hess-21-](https://doi.org/10.5194/hess-21-3859-2017)
- 655 3859-2017, 2017.
- DEFRA: Hydrology Data Explorer, <https://environment.data.gov.uk/hydrology/explore>, 2024.
- Derek Finnie Associates: South Wealden and Eastbourne Dynamic Flood Risk Management (internal), Wokingham, UK, 2022.
- Eggimann, S., Mutzner, L., Wani, O., Schneider, M. Y., Spuhler, D., Moy de Vitry, M., Beutler, P., and Maurer, M.: The Potential of Knowing More: A Review of Data-Driven Urban Water Management, *Environ. Sci. Technol.*, 51, 2538–2553, <https://doi.org/10.1021/acs.est.6b04267>, 2017.
- 660 Ellis, J. B. and Revitt, D. M.: Sewer losses and interactions with groundwater quality, *Water Science and Technology*, 45, 195–202, <https://doi.org/10.2166/wst.2002.0079>, 2002.
- Fileni, F., Fowler, H. J., Lewis, E., McLay, F., Coxon, G., Archer, D., Bruce, E., Yang, L., Fry, M., Cooper, H., and Swain, O.: UK-Flow15-QC: A quality control framework for better river flow data in hydrological research, <https://doi.org/10.5194/egusphere-2026-277>, 7 April 2026.
- 665 Fletcher, T. D., Andrieu, H., and Hamel, P.: Understanding, management and modelling of urban hydrology and its consequences for receiving waters: A state of the art, *Adv. Water Resour.*, 51, 261–279, <https://doi.org/10.1016/j.advwatres.2012.09.001>, 2013.
- Frantzeskaki, N.: Seven lessons for planning nature-based solutions in cities, *Environ. Sci. Policy*, 93, 101–111, <https://doi.org/10.1016/j.envsci.2018.12.033>, 2019.
- 670



- Ge, J., Li, J., Qiu, R., Shi, T., Huang, Z., Liu, Y., and Yuan, Z.: Identifying periods impacted by sewer inflow and infiltration using time series anomaly detection, *Water Res. X*, 25, 100278, <https://doi.org/10.1016/j.wroa.2024.100278>, 2024.
- El Ghazouli, K., El Khatabi, J., Soulhi, A., and Shahrour, I.: Model predictive control based on artificial intelligence and EPA-SWMM model to reduce CSOs impacts in sewer systems, *Water Science & Technology*, 85, 398–408, <https://doi.org/10.2166/wst.2021.511>, 2022.
- Hauser, M., Kelvin, J., and Kleidorfer, M.: Potential of mitigation strategies to compensate climate change impacts on the urban drainage network and pluvial flood hazard, *Water Science & Technology*, 93, 700–718, <https://doi.org/10.2166/wst.2026.237>, 2026.
- Hoard, C. J., Haefner, R. J., Shuster, W. D., Pieschek, R. L., and Beeler, S.: Full Water-Cycle Monitoring in an Urban Catchment Reveals Unexpected Water Transfers (Detroit MI, USA), *Journal of the American Water Resources Association*, 56, 82–99, <https://doi.org/10.1111/1752-1688.12814>, 2020.
- Hosseinzadehtalaei, P., Ishadi, N. K., Tabari, H., and Willems, P.: Climate change impact assessment on pluvial flooding using a distribution-based bias correction of regional climate model simulations, *J. Hydrol. (Amst.)*, 598, 126239, <https://doi.org/10.1016/j.jhydrol.2021.126239>, 2021.
- Ibrahim, Y. A.: Real-Time Control Algorithm for Enhancing Operation of Network of Stormwater Management Facilities, *J. Hydrol. Eng.*, 25, [https://doi.org/10.1061/\(ASCE\)HE.1943-5584.0001881](https://doi.org/10.1061/(ASCE)HE.1943-5584.0001881), 2020.
- Jang, B.-J. and Jung, I.: Development of High-Precision Urban Flood-Monitoring Technology for Sustainable Smart Cities, *Sensors*, 23, 9167, <https://doi.org/10.3390/s23229167>, 2023.
- Kändler, N., Annus, I., Vassiljev, A., and Puust, R.: Real time controlled sustainable urban drainage systems in dense urban areas, *Journal of Water Supply Research and Technology-Aqua*, 69, 238–247, <https://doi.org/10.2166/aqua.2019.083>, 2020.
- Karpf, C., Hoeft, S., Scheffer, C., Fuchs, L., and Krebs, P.: Groundwater infiltration, surface water inflow and sewerage exfiltration considering hydrodynamic conditions in sewer systems, *Water Science and Technology*, 63, 1841–1848, <https://doi.org/10.2166/wst.2011.388>, 2011.
- Kuhaneswaran, B., Sorwar, G., Alaei, A. R., and Tong, F.: Evolution of Data-Driven Flood Forecasting: Trends, Technologies, and Gaps—A Systematic Mapping Study, *Water (Basel)*, 17, 2281, <https://doi.org/10.3390/w17152281>, 2025.
- Li, J., Li, J., Li, Z., Cui, Z., Jing, J., Zou, Y., Li, X., and Xiong, F.: Characteristics and risk assessment of thermal pollution from stormwater runoff in urban lake, *J. Hydrol. (Amst.)*, 658, 133198, <https://doi.org/10.1016/j.jhydrol.2025.133198>, 2025a.
- Li, R., Liu, J., Sun, T., Shao, J., Tian, F., and Ni, G.: Enhancing urban pluvial flood modeling through graph reconstruction of incomplete sewer networks, *Hydrol. Earth Syst. Sci.*, 29, 5677–5694, <https://doi.org/10.5194/hess-29-5677-2025>, 2025b.
- Macdonald, N., Redfern, T., Miller, J., and Kjeldsen, T. R.: Understanding the impact of the built environment mosaic on rainfall-runoff behaviour, *J. Hydrol. (Amst.)*, 604, 127147, <https://doi.org/10.1016/j.jhydrol.2021.127147>, 2022.
- Mallakpour, I. and Villarini, G.: Analysis of changes in the magnitude, frequency, and seasonality of heavy precipitation over the contiguous USA, *Theor. Appl. Climatol.*, 130, 345–363, <https://doi.org/10.1007/s00704-016-1881-z>, 2017.



- Met Office: Eastbourne Location-specific long-term averages, [https://www.metoffice.gov.uk/research/climate/maps-and-](https://www.metoffice.gov.uk/research/climate/maps-and-data/location-specific-long-term-averages/u100y2rs8)
 705 [data/location-specific-long-term-averages/u100y2rs8](https://www.metoffice.gov.uk/research/climate/maps-and-data/location-specific-long-term-averages/u100y2rs8), 2026.
- Mullapudi, A., Wong, B. P., and Kerkez, B.: Emerging investigators series: building a theory for smart stormwater systems, *Environmental Science-Water Research & Technology*, 3, 66–77, <https://doi.org/10.1039/c6ew00211k>, 2017.
- Mullapudi, A., Bartos, M., Wong, B., and Kerkez, B.: Shaping Streamflow Using a Real-Time Stormwater Control Network, *SENSORS*, 18, <https://doi.org/10.3390/s18072259>, 2018.
- 710 Mullapudi, A., Lewis, M. J., Gruden, C. L., and Kerkez, B.: Deep reinforcement learning for the real time control of stormwater systems, *Adv. Water Resour.*, 140, <https://doi.org/10.1016/j.advwatres.2020.103600>, 2020.
- Najafi, H., Shrestha, P. K., Rakovec, O., Apel, H., Vorogushyn, S., Kumar, R., Thober, S., Merz, B., and Samaniego, L.: High-resolution impact-based early warning system for riverine flooding, *Nat. Commun.*, 15, 3726, [https://doi.org/10.1038/s41467-](https://doi.org/10.1038/s41467-024-48065-y)
 024-48065-y, 2024.
- 715 Naughton, J., Sharior, S., Parolari, A., Strifling, D., and McDonald, W.: Barriers to Real-Time Control of Stormwater Systems, *J. Sustain. Water Built Environ.*, 7, <https://doi.org/10.1061/JSWBAY.0000961>, 2021.
- Nedergaard Pedersen, A., Wied Pedersen, J., Viguera-Rodriguez, A., Brink-Kjær, A., Borup, M., and Steen Mikkelsen, P.: The Bellinge data set: open data and models for community-wide urban drainage systems research, *Earth Syst. Sci. Data*, 13, 4779–4798, <https://doi.org/10.5194/essd-13-4779-2021>, 2021.
- 720 Oh, J. and Bartos, M.: Scalable, adaptive and risk-informed design of hydrological sensor networks, *Nature Water*, 3, 1144–1154, <https://doi.org/10.1038/s44221-025-00496-7>, 2025.
- Qiu, C., Shao, G., Zhang, Z., Zhou, C., Hou, Y., Zhao, E., Guo, X., and Guan, X.: Unsupervised Real Time and Early Anomalies Detection Method for Sewer Networks Systems, *IEEE Access*, 12, 21698–21709, <https://doi.org/10.1109/ACCESS.2024.3359302>, 2024.
- 725 Rahman, M., Smith, V., Kremer, P., Wadzuk, B., Jiao, X., Amur, A., and Arabi, S.: Incorporating artificial intelligence into the future of stormwater management, *Discover Applied Sciences*, 8, 524, <https://doi.org/10.1007/s42452-026-08488-2>, 2026.
- Ramovha, N., Chadyiwa, M., Ntuli, F., and Sithole, T.: The Potential of Stormwater Management Strategies and Artificial Intelligence Modeling Tools to Improve Water Quality: A Review, *Water Resources Management*, 38, 3527–3560, <https://doi.org/10.1007/s11269-024-03841-9>, 2024.
- 730 Roosipuu, P., Annus, I., Kuusik, A., Kändler, N., and Alam, M. M.: Monitoring and control of smart urban drainage systems using NB-IoT cellular sensor networks, *Water Science & Technology*, 88, 339–354, <https://doi.org/10.2166/wst.2023.222>, 2023.
- Roy, A. H., Wenger, S. J., Fletcher, T. D., Walsh, C. J., Ladson, A. R., Shuster, W. D., Thurston, H. W., and Brown, R. R.: Impediments and Solutions to Sustainable, Watershed-Scale Urban Stormwater Management: Lessons from Australia and the
 735 United States, *Environ. Manage.*, 42, 344–359, <https://doi.org/10.1007/s00267-008-9119-1>, 2008.
- Santos, L. B. L., Escobar-Silva, E. V., Satolo, L. F., Oyarzabal, R. S., Diniz, M. M., Negri, R. G., Lima, G. R. T., Stephany, S., Soares, J. A. J. P., Duque, J. S., Saraiva Filho, F. L., Caseri, A. N., and Bacelar, L.: Machine learning-based hydrological



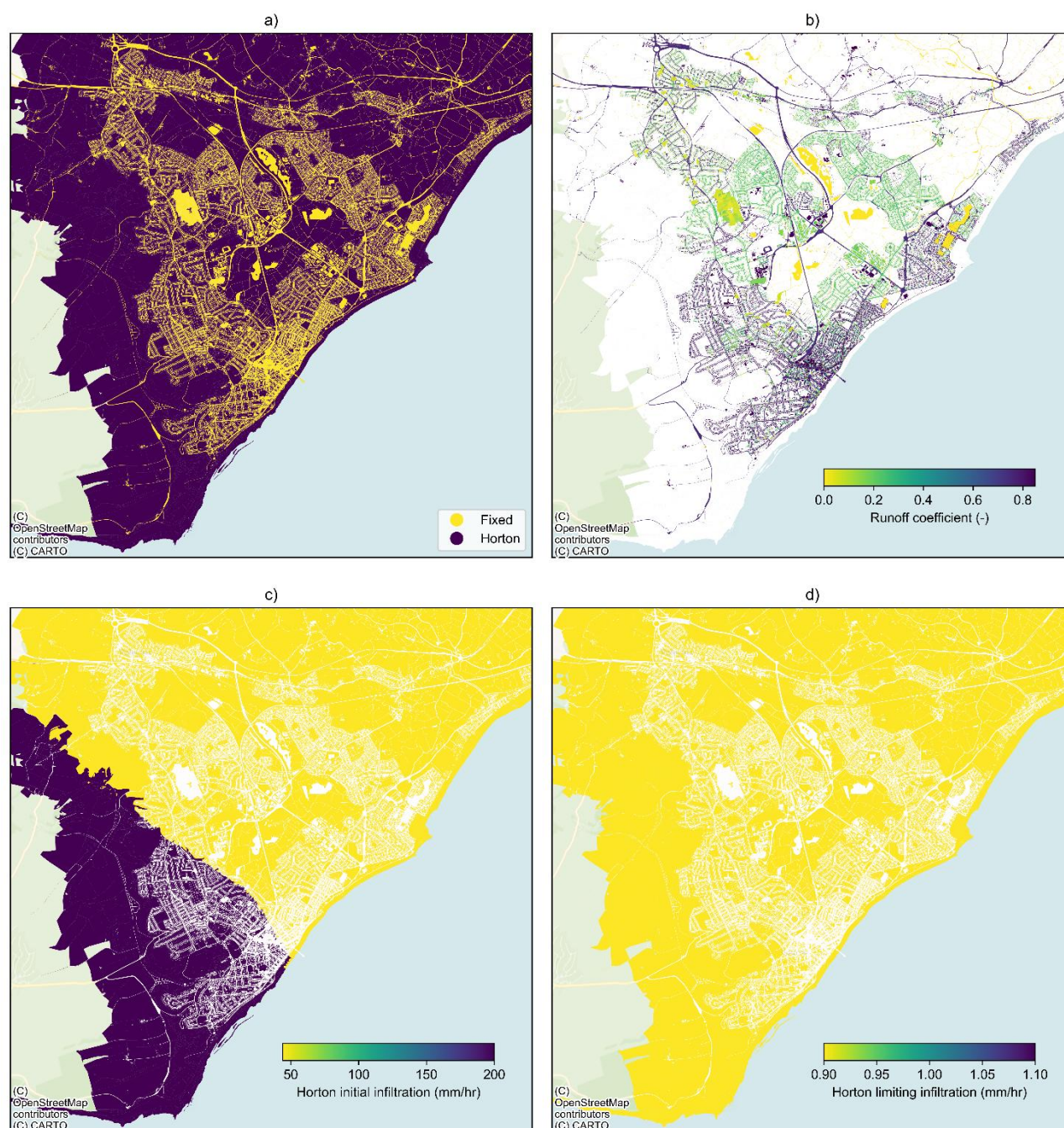
- models for flash floods: a systematic literature review, *Smart Construction and Sustainable Cities*, 3, 21, <https://doi.org/10.1007/s44268-025-00071-9>, 2025.
- 740 Schmitt, Z. K., Hodges, C. C., and Dymond, R. L.: Simulation and assessment of long-term stormwater basin performance under real-time control retrofits, *Urban Water J.*, 17, 467–480, <https://doi.org/10.1080/1573062X.2020.1764062>, 2020.
- Sharifi, A., Tarlani Beris, A., Sharifzadeh Javidi, A., Nouri, M., Gholizadeh Lonbar, A., and Ahmadi, M.: Application of artificial intelligence in digital twin models for stormwater infrastructure systems in smart cities, *Advanced Engineering Informatics*, 61, 102485, <https://doi.org/10.1016/j.aei.2024.102485>, 2024.
- 745 Sharior, S., McDonald, W., and Parolari, A. J.: Improved reliability of stormwater detention basin performance through water quality data-informed real-time control, *J. Hydrol. (Amst.)*, 573, 422–431, <https://doi.org/10.1016/j.jhydrol.2019.03.012>, 2019.
- Southern Water: Water unveils industry-leading £15 million Smart Sewer Technology, <https://www.southernwater.co.uk/latest-news/water-unveils-industry-leading-15-million-smart-sewer-technology-in-battle-against-pollution/>, 2024.
- 750 Subehi, L., Fukushima, T., Onda, Y., Mizugaki, S., Gomi, T., Kosugi, K., Hiramatsu, S., Kitahara, H., Kuraji, K., and Terajima, T.: Analysis of stream water temperature changes during rainfall events in forested watersheds, *Limnology (Tokyo)*, 11, 115–124, <https://doi.org/10.1007/s10201-009-0296-2>, 2010.
- Sweetapple, C., Webber, J., Hastings, A., and Melville-Shreeve, P.: Realising smarter stormwater management: A review of the barriers and a roadmap for real world application, <https://doi.org/10.1016/j.watres.2023.120505>, 2023.
- 755 Sweetapple, C., Illingworth, T., and Melville-Shreeve, P.: Dataset: High resolution hydrometric and sewer system monitoring of an urban catchment with complex flood risk, <https://doi.org/https://doi.org/10.5281/zenodo.20699634>, 2026a.
- Sweetapple, C., Illingworth, T., and Melville-Shreeve, P.: Computer software: Blue Heart dataset scripts: v1.0.0, <https://doi.org/https://doi.org/10.5281/zenodo.21390823>, 2026b.
- Wang, M., Zhang, M., Shi, H., Huang, X., and Liu, Y.: Uncertainty analysis of a pollutant-hydrograph model in assessing inflow and infiltration of sanitary sewer systems, *J. Hydrol. (Amst.)*, 574, 64–74, <https://doi.org/10.1016/j.jhydrol.2019.04.011>, 2019.
- 760 Wang, Y., Huang, B., and Zhu, D. Z.: Assessment of rainfall-derived inflow and infiltration in sewer systems with machine learning approaches, *Water Science & Technology*, 89, 1928–1945, <https://doi.org/10.2166/wst.2024.115>, 2024.
- Wartalska, K., Szymczewski, S., Domalewska, W., Wdowikowski, M., Przestrzelska, K., Kotowski, A., and Kaźmierczak, B.: The Impact of Climate Change on the Functioning of Drainage Systems in Industrial Areas—A Case Study, *Atmosphere (Basel)*, 16, 347, <https://doi.org/10.3390/atmos16030347>, 2025.
- 765 Webber, J. L., Fletcher, T., Farmani, R., Butler, D., and Melville-Shreeve, P.: Moving to a future of smart stormwater management: A review and framework for terminology, research, and future perspectives, <https://doi.org/10.1016/j.watres.2022.118409>, 2022.
- 770 Xu, W. D., Burns, M. J., Cherqui, F., and Fletcher, T. D.: Enhancing stormwater control measures using real-time control technology: a review, *Urban Water J.*, 18, 101–114, <https://doi.org/10.1080/1573062X.2020.1857797>, 2021.



- Zhang, H., Xu, Z., Wang, W., Peng, S., Li, C., Fang, S., Guo, D., and Yin, H.: Quantifying the performance of urban sewer network using inverse-problem models: An approach for synchronous determination of in-sewer groundwater infiltration and pollutant degradation, *Journal of Hydrodynamics*, 37, 1–13, <https://doi.org/10.1007/s42241-025-0002-9>, 2025.
- 775 Zheng, Y., Chen, W., Xu, C., Zhang, Y., Jin, X., Zhou, Y., and Farmani, R.: A knowledge-based deep learning model for accurate urban drainage system prediction under spatiotemporally variable rainfall, *Water Res.*, 300, 125930, <https://doi.org/10.1016/j.watres.2026.125930>, 2026.



780 Appendix A: Catchment infiltration and runoff characteristics



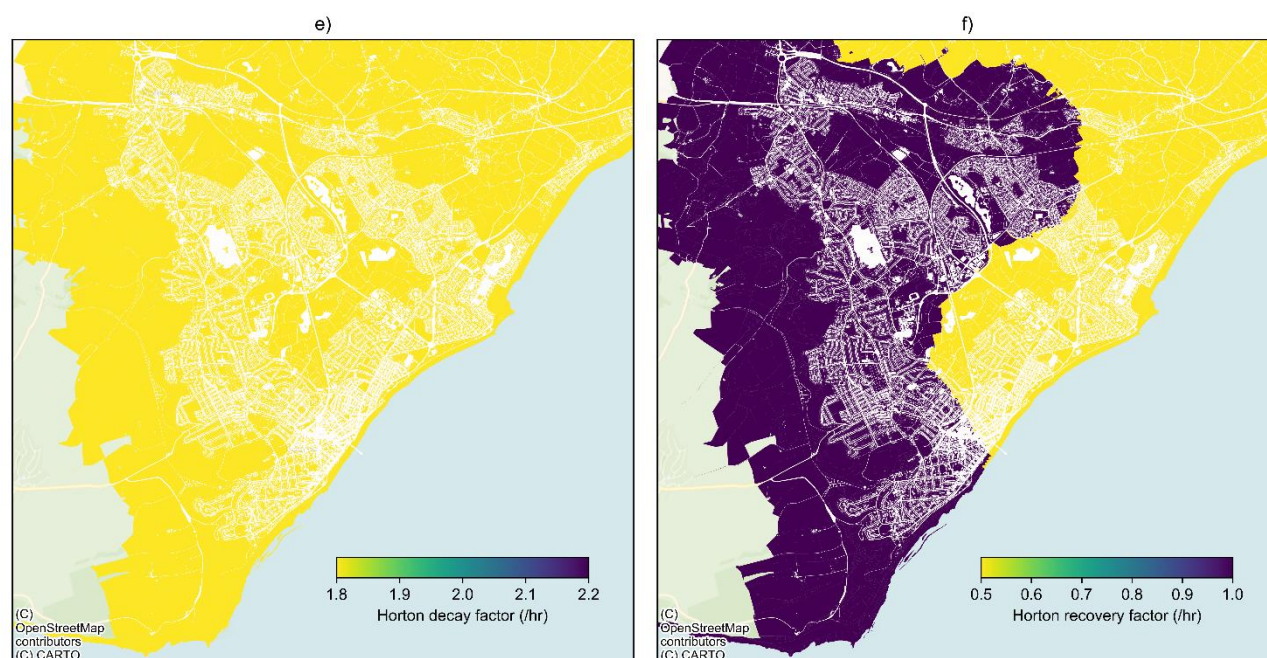


Figure A1: Runoff characteristics extracted from the hydraulic model of the case study area (discussed in Section 1): a) Areas in which different runoff models area applied (fixed coefficient or Horton); b) Catchment runoff coefficient values for areas modelled with fixed coefficient; c) Initial infiltration values for areas using Horton model; d) Recovery factor values for areas using Horton model; e) Limiting infiltration values for areas using Horton model; e) Decay factor values for areas using Horton model.