https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

UK-Flow15-QC: A quality control framework for better river flow data in hydrological research

Felipe Fileni[1], Hayley J. Fowler[1], Elizabeth Lewis[1,2], Fiona McLay[3,4], Gemma Coxon[5], David Archer[1,6], Emma Bruce[3], Longzhi Yang[7], Matt Fry[8], Hollie Cooper[8], Ollie Swain[8]

[1] School of Engineering, Newcastle University, Newcastle upon Tyne, UK
[2] Department of Mechanical, Aerospace and Civil Engineering, University of Manchester, Manchester, UK
[3] Scottish Environment Protection Agency (SEPA), Motherwell, UK
[4] Environment Agency, Tyneside House, Newcastle upon Tyne, UK
[5] School of Geographical Sciences, University of Bristol, Bristol, UK
[6] JBA Trust, North Yorks, UK
[7] Department of Computer and Information Sciences, Northumbria University, Newcastle Upon Tyne, UK
[8] UK Centre for Ecology & Hydrology, Maclean Building, Crowmarsh Gifford, Wallingford, UK

Correspondence to: Felipe Fileni (F.De-Mendonca-Fileni2@newcastle.ac.uk)

Abstract. The significant increase in computing power over the past 70 years has progressively enabled the use of extensive datasets for hydrological modelling. The colossal scale of these datasets, i.e., over one million timesteps per station for a 30-year record at 15-min resolution, makes implementing effective quality control (QC) particularly challenging. In this study, we present a national-scale, open-source quality-control framework tailored for the UK's 15-minute river flow dataset, UK-Flow15, which is described in Part 1 of this paper series. The framework combines manual visual inspection of anomalies with automated detection of statistical artefacts, incorporating both established and novel procedures. In particular, we introduce methods to evaluate high-flow events by comparing them with rainfall records and flow observations from neighbouring catchments. Application of the framework within a UK dataset reveals that while many stations maintain generally reliable records, over 20% exhibit visually identifiable issues such as truncations, discontinuities, or missing data. Automated checks indicate that most (78%) stations contain at least isolated segments of suspicious behaviour. Our high-flow event validation procedures confirm most peak flows, but also flag a small proportion of events as potentially spurious due to a lack of consistency with nearby flow (10.5%) or rainfall (14.5%) support. We further demonstrate that data quality has a measurable impact on hydrological modelling, with catchments containing flagged anomalies producing the least reliable simulations in terms of NSE and High-Flow Bias. By making flagged data and metadata openly accessible, the framework enables users to make informed decisions about data suitability. This work highlights the critical importance of rigorous QC in sub-daily hydrology and provides a scalable tool to support the development of more reliable, high-resolution hydrological data.

1. **Introduction**

Moore's Law, proposed by Gordon Moore in 1965, describes the trend that computing power, as measured by transistor density, doubles roughly every two years. This exponential growth has enabled significant progress in the Earth Sciences,

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

where both data and models have expanded in complexity and resolution. For instance, climate model resolutions have advanced from the coarse scales of the 1990s (300 km², monthly; Hulme & Jenkins, 1998) to the high-resolution UKCP18-local dataset (2.2 km², hourly; L. Kendon et al., 2023) showing how technology can shape scientific tools.

Flood events highlight the importance of these high-resolution simulations in Earth sciences. Reports from as early as the 1500s describe river level rises of several meters within hours, leading to severe impacts (Archer and Fowler, 2021; Champion, 1861). Today, science shows that extreme hydrometeorological events require high temporal resolution data to be properly characterised (Hou et al., 2020; Huang et al., 2019). Without this resolution, key features of rapid-onset events may be missed in models, reducing their effectiveness in risk assessment (Fileni et al., 2025a).

With current computational capabilities supporting the production and analysis of detailed data, and with a growing emphasis on data-driven decision-making and disaster management, high-resolution rainfall data, i.e. large-scale sub-daily rainfall datasets, have become increasingly accessible. Tools such as radar, which typically record data at sub-hourly timescales, have enabled rapid advancements in flash flood nowcasting (Costabile et al., 2023; Green et al., 2025; Heuvelink et al., 2020; Li et al., 2024). Gridded sub-daily datasets, such as CEH-Gear-1hr (Lewis E. et al., 2022) coupled with downscaled local climate projections (Chan et al., 2018; Kendon et al., 2014), have facilitated the development of policies for climate change adaptation for extreme rainfall events (Chan et al., 2023).

Adaptation to extreme flows, however, has lagged, partially due to the scarcity of large-scale sub-daily flow datasets. This absence is noted within studies focusing on sub-daily flows, that are often calibrated using daily data (Bauwe et al., 2017; Campbell et al., 2018; Wetterhall et al., 2011; Yang et al., 2016). Still, sub-daily data is essential for accurately modelling flood waves, as daily data tend to systematically underestimate flood peaks (Yu et al., 2018). Accurate model calibration with high resolution flows is especially important when accounting for climate change, as error propagation due to coarser temporal resolutions can cause further underestimations of flood risk (Kim et al., 2018). This can be especially important for estimating large return period events where flows are highly sensitive to temporal storm distributions (Zhou et al., 2021).

Historical sub-daily flow records are available for some locations worldwide, but these datasets are scattered across national Application Programming Interfaces (APIs) without standardization. For instance, in the US, data is typically available at a 5-min resolution, generally recorded in cubic feet per second (USGS Surface-Water Data for the Nation); in France, records are available in resolutions ranging from 5 to 60 mins (Hydrométrie); and in Sweden, data are recorded every 15 mins, typically in cubic meters per second (Ladda ner observationer från sjöar och vattendrag). Although these APIs represent significant progress in data accessibility, they are still often constrained by limitations such as maximum request sizes or interfaces that are difficult to navigate.

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

Beyond accessibility, data quality and consistency remain major challenges. Flow records are subject to numerous sources of uncertainty, from equipment malfunctions and data management issues to limitations in rating curves and extrapolation methods (Di Baldassarre and Montanari, 2009; Coxon et al., 2015; Mcmillan et al., 2012; Mcmillan and Westerberg, 2015). While many methods exist to characterize and reduce rating curve uncertainty at higher temporal resolutions (McMahon et al., 2025; Tomkins, 2014), they do not focus on addressing basic data quality issues, such as the malfunctioning of sensors or technical failures.

Some established routines exist for the basic quality control of flow data, but they are primarily designed for daily resolution and require adaptation for finer timescales. These routines can detect issues such as repeated values and statistical anomalies effectively at the daily level (Crochemore et al., 2020; Gudmundsson et al., 2018). In contrast, tailored quality control systems for high-resolution rainfall and hydrological data have demonstrated the value of such approaches in identifying spurious data (Lewis et al., 2021; Schmidt et al., 2023).

The UK holds extensive sub-daily flow records, commonly digitized at 15-min intervals. Despite their availability, these records were underused, largely due to limited computational resources and a lack of demand at the time they were collected. Consequently, efforts focused primarily on daily summary statistics, such as records of the max, mean and min daily values. Sub-daily data is rarely analysed or quality-assured. With current advances in computing and data accessibility, there is now an opportunity, and a need, to revisit the neglected data. However, their historical underuse means that many potential issues remain unaddressed. A systematic, national-scale assessment is essential to evaluate and improve the quality of this high-resolution archive for reliable hydrological applications (Fileni et al., 2023).

In this paper, we present a quality-control framework developed for, and applied to, the UK's national 15-minute river flow dataset (Fileni et al., 2025b). The framework was implemented across the full dataset prior to release, and the resulting QC flags and metadata form part of the published archive. Initially, we outline the rationale and ideas used for developing and tailoring this framework for our current data. Next, we introduce methods derived from our research to address the basic quality issues within the dataset. Finally, we illustrate potential outcomes and discuss the advantages of a quality-controlled, high-resolution UK flow dataset for the broader hydrological community.

2. Development of a Quality Control framework

In the UK, nationally consistent and QC'd river flow data is available through the National River Flow Archive (NRFA), but only at daily resolution or for annual flood statistics. The available sub-daily hydrometric data is characterised by varying quality standards, inconsistent metadata, and even data at varying resolution, which poses an obstacle in generating a coherent, high-resolution dataset. The many versions and updates to rating curves, along with corrections that have often been applied

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

only to peak flow datasets and not reflected in the original continuous flow records, have introduced inconsistencies in UK flow data. These inconsistencies have been made more difficult to track due to the lack of version control in the original 15-min version of the data. In our QC framework, we aim to clarify these inconsistencies and to create a sub-daily flow dataset that is as standardized and as uniform as possible. In developing this framework, we revisited traditional concepts of data curation and governance, which will be further described in this section.

To maximise transparency and preserve the original information content of the records, UK-Flow15 dataset was assembled with minimal intervention. Most observations were retained exactly as provided by the data sources. Limited preprocessing was applied to harmonise temporal resolution and resolve duplicated timestamps with conflicting values, following clearly defined and fully documented procedures. These steps produced a consistent 15-minute input dataset for the quality-control framework described in this paper. All such steps were applied systematically prior to the quality-control procedures described here and were fully documented as part of the dataset preparation. Despite these measures, the remaining time series data still contain significant quality issues and notable variability in quality across stations. Releasing the dataset in its initial state, lacking robust quality assurance or validation mechanisms, was incompatible with current standards of big data management and principles (Wilkinson et al., 2016). Consequently, we developed a structured framework for QC, enabling thorough auditing and evaluation by future users.

Upon examining the data, multiple quality issues were identified. Certain stations exhibited clear and visually detectable anomalies, such as spurious or incomprehensible values, assessed initially through visual plotting. Further comparisons with existing UK hydrological products revealed discrepancies, for example, between published daily flow data and the averages computed from the corresponding 15-min data, or mismatches between reported flood statistics - annual maxima (AMAX) or peaks-over-threshold (POT) records and those derived directly from sub-daily data. To maintain transparency, these discrepancies were preserved in the dataset, accompanied by detailed metadata to enable verification in scenarios where accuracy is critical.

Given the necessity for robust quality checks and the absence of existing methodologies specifically tailored for high-resolution flow data, we adapted quality control techniques from established daily flow quality assurance and sub-daily rainfall quality assurance procedures (Blenkinsop et al., 2017; Crochemore et al., 2020; Gudmundsson et al., 2018; Lewis et al., 2018a, 2021; Schmidt et al., 2023). These adapted methods aim to identify and flag anomalies within the dataset, recognizing the significant challenge posed by the extensive diversity and variability inherent of a national-scale dataset. Setting thresholds for anomaly detection thus required testing to identify suitable universal thresholds. Furthermore, recognizing that the primary application of high temporal resolution flow data is in assessing high flow events, we developed novel verification methods for high flow accuracy. These methods use not only catchment-specific data but also supplementary datasets such as rainfall observations and flows from neighbouring catchments.

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

Finally, to ensure the quality and integrity of the dataset, all scripts and algorithms utilized in performing these checks have been made publicly available, inviting external verification and scrutiny.

3 Quality control procedures

3.1 Visual inspection anomalies

An initial visual assessment was conducted to evaluate the overall quality of stations within the dataset. This process involved plotting the monthly mean, maxima, and standard deviation for each station and visually inspecting the time series for anomalies, with any unusual behaviour being recorded in a spreadsheet. After this first initial assessment, the stations were reviewed, aiming to classify the issues and facilitate the user to identify these issues and understand their potential impact on analysis of the dataset. We have classified these issues into three categories; scope, type and confidence level (Figure 1).

The first classification, "scope", specifies the duration of time that the issue is visible and is divided into two categories:

1. full_timeseries, stations in which the entire time series exhibited substantial issues that could compromise scientific reliability (Figure 1.a – Station 31022);
2. partial_timeseries, stations with problematic segments or discontinuous records, exemplified by station 30012, truncated on the highlighted section in light red (Figure 1a – Station 30012).

The second classification, "type", specifies the type of issue and is divided into four categories:

1. continuity issues, stations exhibiting abrupt shifts in flow values that look visually unnatural/caused by external factors (Figure 1b – Station 28015);
2. truncation, instances where maximum flow values appear to be artificially capped at a specific limit, with repetition of the same maximum value across several steps of the timeseries (Figure 1b – Station 41013);
3. traces of data, where initial data in the timeseries is disjointed from the main timeseries archive (Figure 1b – Station 22003);
4. short timeseries, a station where there is less than 5 years of data or with only short data fragments (Figure 1b – Station 69011).

Additionally, some anomalies did not fit into the above categories and were labelled as other data issues (Figure 1b – Station 39852).

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

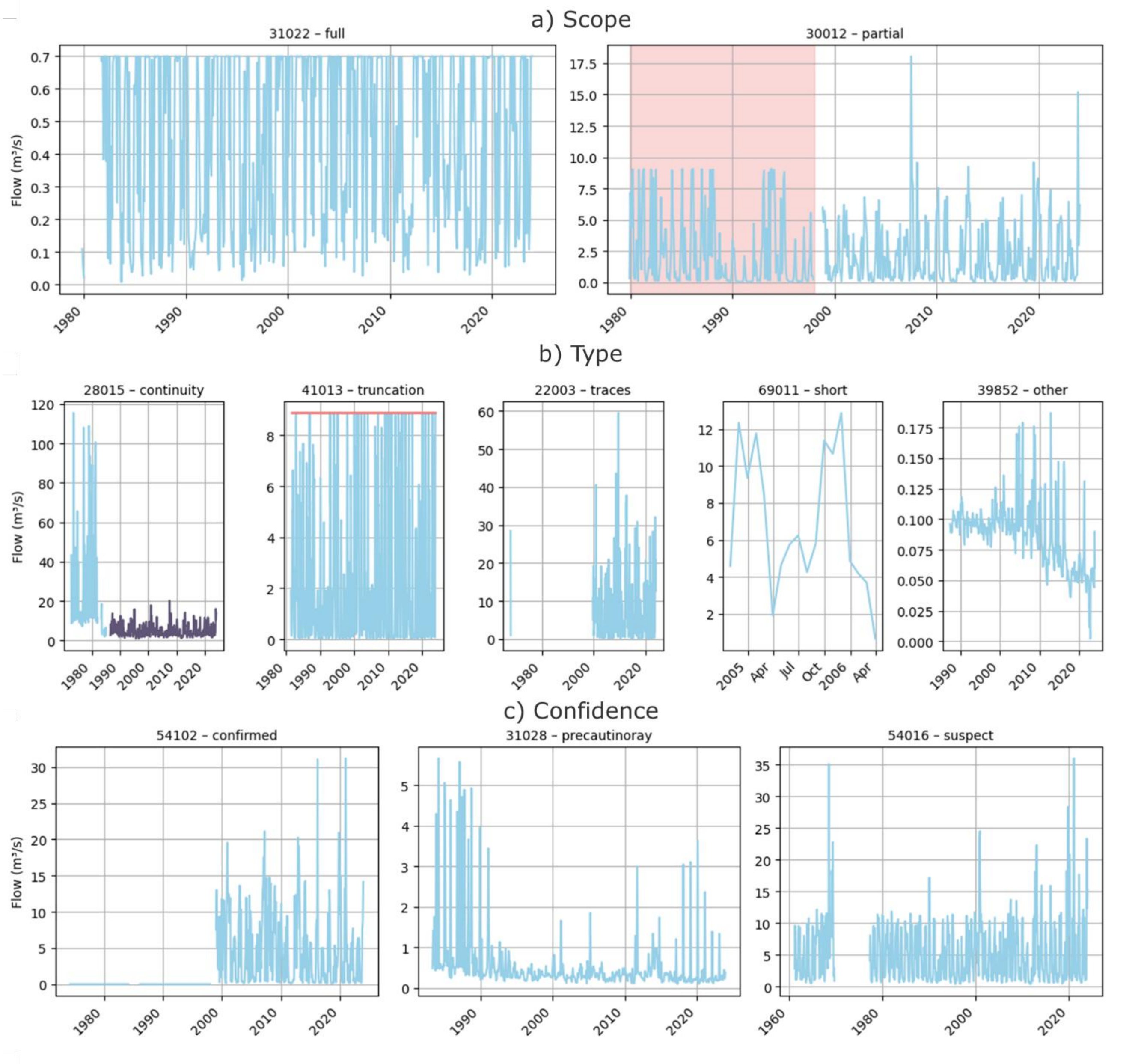


Figure 1: Classification of common anomalies by a) scope, b) type and c) confidence, with examples given at different flow stations.

Finally, the third classification, "confidence", specifies the severity of the issue and is divided into three categories:

165 1. confirmed, there is an issue at the station that is expected to affect scientific analysis, and needs to be reviewed before use., For instance, station 54102 that has maximum values that are very close to 0 for a very large portion of the record (Figure 1c – Station 54102);

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

2. precautionary, the issue might affect scientific analysis and should be reviewed if the analysis requires a higher degree of accuracy/more rigorous curation. As in station 31028 where the series before 1992 presents maximum flows that are considerably different than the values recorded after (Figure 1.c -Station 31028);
3. suspect, the issue is unlikely to significantly influence the analysis and should only be excluded if very high-quality data are required. It may also serve as a point of reference if issues are encountered when using the station in scientific analysis. For instance station 54016 in which the period of data from 1960 to 1970 present slightly anomalous behaviour when compared to the rest of the timeseries (Figure 1c – Station 54016)

3.2 Verification with other UK products

Hydrological analysis in the UK relies on two derived datasets based on 15-min records from the National River Flow Archive (NRFA). The national daily flow records – NRFA daily - are an average of the 15-min records from 09:00 on one day until 08:45 the next day. The peak flow statistics – NRFA AMAX and POT - are, respectively, the annual maximum instantaneous recorded flow, and the flows above an established threshold which aims to average 5 events per year for record duration. Despite being derived from 15-min data, these records sometimes diverge due to rating curve changes, version control inconsistencies, and station-specific issues. Discrepancies between the NRFA AMAX and POT values may be particularly common, as peak flows undergo annual review and contain corrections that may not be reflected in the original 15-min datasets. These reviews aim to manually identify potential issues affecting peak flows, with a focus on stations that are a cornerstone of flood risk methodologies in the UK. These stations provide the underlying data for statistical methods used in the estimation of extreme events (Robson and Reed, 1999). Historically, 15-minute flow data have not played a role in these applications, and as a result, corrections made during peak flow reviews have not always been propagated back to the original high-resolution datasets.

When a timestep was available in both series, the values were compared. The comparison between NRFA daily and the 15-min series was done by averaging the 15-min records within the hydrological day, whilst the comparison between NRFA POT and AMAX was done on a timestamp basis.

3.3 Traditional quality control

The traditional quality control procedures aimed to identify suspicious data, defined as measurements that are physically unrealistic or inconsistent, such as negative flows or abrupt changes in recorded values. We identified and categorized eight types of issues as shown in Figure 2 and described as Equations in Table 1.

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

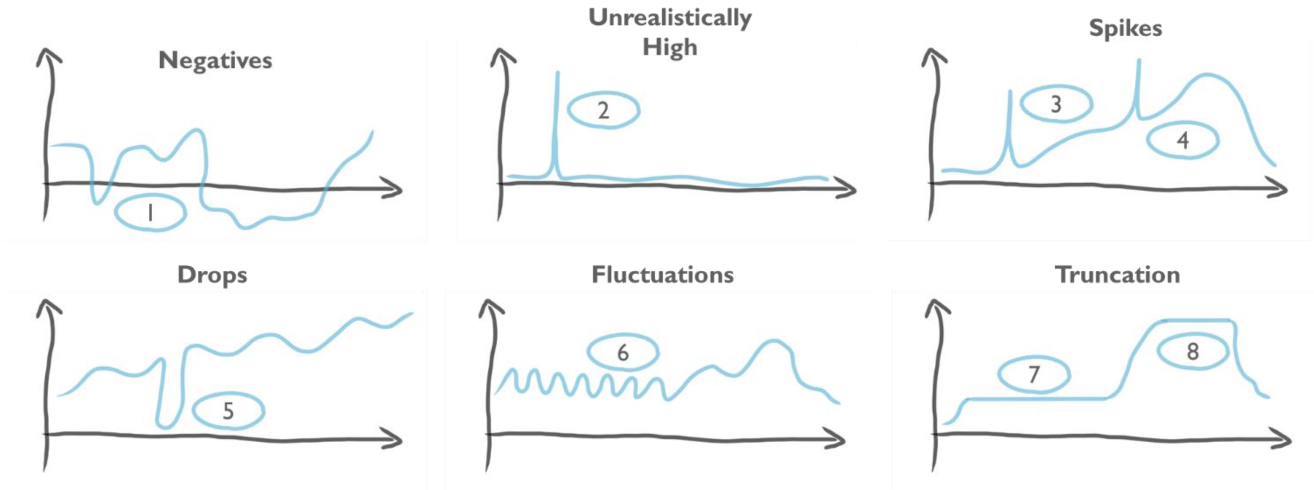


Figure 2: Representation of eight common artefacts that can be found in UK flow data .

Table 1: Mathematical representation of each of the traditional quality control issues computed

Nº	Quality control issue	Equation(s)
1	Negative	$q_x < 0$
2	Unrealistic Flow	$q_x > 5000m^3/s$
3	Relative Spike	$(q_x > 5q_{x-1} \wedge q_x < 5q_{x+1}) \vee (q_x > \frac{q_{x-1}}{5} \wedge \ q_x < \frac{q_{x+1}}{5})$
4	Absolute Spike	$(q_x - q_{x-1} > Q99 \wedge q_x - q_{x+1} > Q99) \vee (q_{x-1} - q_x > Q99 \wedge q_{x+1} - q_x > Q99)$
5	Drop	$q_x < 5q_{x+1}$
6	Fluctuations	$Sign\ check: s_i = sign(q_i - q_{i-1}) == \{-1,0,1\}$ $One\ oscillation: O_i = s_i * s_{i+1} < 0$ $\sum_{i=x-15}^{x-1} O_i > 8$
7	Low Truncation	$Constant\ value\ check: C_i = if\ (q_i == q_{i-1}) \rightarrow 1$ $\sum_{i=x-671}^{x} C_i == 672$
8	High Truncation	$\sum_{i=x-95}^{x} C_i == 96\ \wedge q_x > Q_{99}$

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

Negative values (i.e. any timestep below 0 m³/s) are flagged as they are physically unrealistic (Table1/Figure 2 – Issue 1); Unrealistically high flows refer to values that are physically impossible for any river covered in this study (Table1/Figure 2 – Issue 2). The highest validated peak flow in the UK was recorded in the River Tay at just under 2500 m³/s. To ensure consistency with the EA and SEPA forecasting systems, we set the threshold for unrealistically high flows at 5000 m³/s.

"Spikes" consist of rapid flow increases followed by an instant flow decrease, often caused by faulty instrumentation. Two metrics were used to categorize spikes: a relative metric (Table1/Figure 2 – Issue 3), that is, a sudden increase and decrease relative to the current flow state; and an absolute metric (Table1/Figure 2 – Issue 4), a pre-defined threshold based on a high flow value for the station timeseries. Relative spikes were flagged when flows increased by 5 times the original value (5x), with a similar decrease in the next timestep. Absolute spikes were identified using the station's Q99 high-flow value: any instance where the flow increased beyond Q99 and then decreased by an equivalent amount was flagged.

Due to absence of literature in this area, the sensitivity of these spike thresholds was tested by varying their magnitudes, specifically at factors of 1x 2x, 5x, and 10x (Figure 3a). A 1x spike (100% increase) frequently occurred, being flagged in almost all stations (median occurrence: 18 times per station), thus identifying many plausible, minor flow variations. Increasing the threshold significantly reduced the number of flagged stations. At the 2x threshold, the number of flagged stations decreased, demonstrating improved discrimination of genuine artefacts. At a 5x threshold, around one-third of stations were flagged, and the median number of flagged timesteps per station stabilized at four. This reduction became less pronounced when further increasing the threshold to 10x. Consequently, the 5x threshold was selected as optimal, balancing sensitivity to genuine anomalies

For absolute spikes, sensitivity analyses were conducted using quantiles Q90, Q95, Q98, and Q99. Low-flow variations frequently triggered flags, especially in perennial rivers, where low flows commonly skew high but not exceptionally high quantile thresholds. Increasing the quantile threshold resulted in fewer flagged stations, with the median number of flagged timesteps decreasing from three at Q90 to one at Q98 and Q99. The Q99 threshold was ultimately selected due to its widespread acceptance in hydrology and good balance between sensitivity and identification of genuine spikes without excessive false positives (Figure 3b).

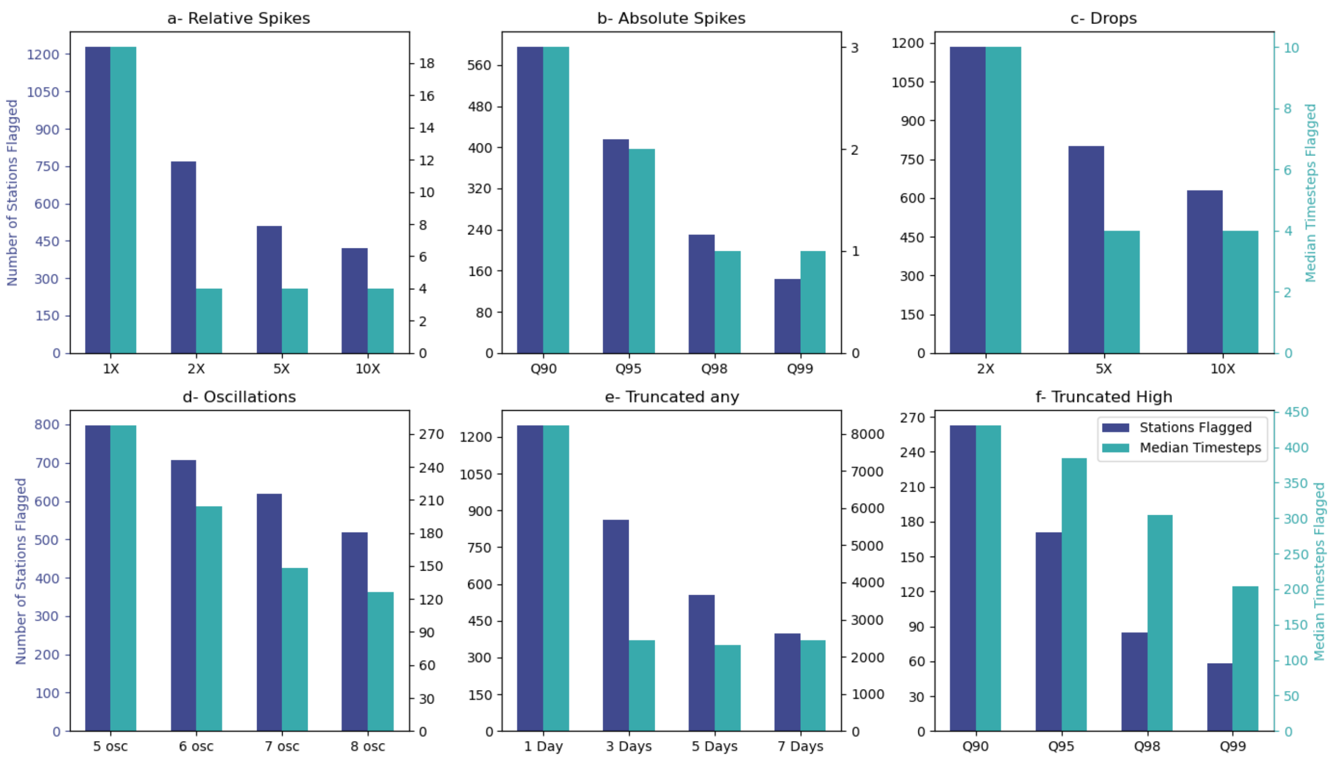
Additionally, rapid drops (Table1/Figure 2 – Issue 5) were flagged as suspicious, as sudden declines in flow over short periods are often unrealistic. Sensitivity testing (Figure 3c) showed a consistent decrease in flagged values as the threshold increased. We adopted the same 5x threshold to keep the QC consistent with the spikes value.

Fluctuations (Table1/Figure 2 – Issue 6) are changes in flow where the first derivative alternates in sign for consecutive timesteps. These are common occurrences at a considerable number of stations, with between 523 to 798 stations flagged,

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

depending on the threshold used (Figure 3d). We have tried to calibrate this parameter to flag stations where this issue was recurrent enough, with a 16 timesteps (4-hours) being the value chosen from visual analysis of reservoir or hydro-electric power influenced stations, the locations with high frequency of fluctuations. The number of fluctuations required for the station to be flagged has been sensitivity-tested on the whole dataset, with a high number of stations flagged for any of the thresholds chosen, we have ended up deciding to pick the more strict threshold, of 8 fluctuations over 16 timesteps (Figure 3.d) as it allows the flagging of the stations where this issue is problematic.

Truncations occur when flow values remain constant for an extended period, often due to instrumentation faults, e.g. weed growth affecting the instrument. Truncations were defined by two thresholds: low-flow truncations (Table1/Figure 2 – Issue 7), which are more common, as stable low flows are plausible and often a consequence of physical processes, such as perennial rivers; and high-flow truncations (Figure 2 – Issue 8), less common, as high flows usually exhibit greater variability. After sensitivity testing (Figure 3e,f), we flagged truncated high flows (Q99+) persisting for more than a day and low flows truncated for over a week. Even though these thresholds have been established aiming to identify spurious data and failures of equipment, we understand that low truncations lasting a week are plausible, notably on river with very low flows occurrences. The same is true for high flow truncation thresholds, with stations that present a high baseflow index presenting plausible truncated flows above Q99. However, increasing this threshold further would hamper the identification of this artifact, hence the choice of thresholds.



https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

Figure 3: Sensitivity testing of the qualitative quality control parameters. Number of stations flagged for different tests for different thresholds: (a) relative spikes; (b) absolute spikes; (c) drops; (d) fluctuations; (e) truncations low; (f) truncations high. X-axis: factor - a factor increase from 1x (doubled) to 10x; percentile – flow percentile used from Q90 to Q99. Y-axis 1: number of stations flagged. Y-axis 2: Median timesteps flagged per station.

3.4 High flows quality control

One of the key objectives of using high-resolution flow data was to accurately capture peak and high-flow events. To assess their quality, we implemented specific checks focusing on extremes. First, we applied methods from previous studies(Crochemore et al., 2020; Gudmundsson et al., 2018), including flagging flow values that abnormally deviated from the logged flow distribution (Table 2 – Issue 1) and annual maximum values that were more than double the second ranked annual maximum recorded at the station (Table 2 – Issue 2).

Next, we fitted a Generalized Extreme Value (GEV) distribution to the annual maxima flow series at each station with at least 10 years of data. The three parameters of the distribution: shape, scale and location of the distribution were estimated using maximum likelihood estimation (MLE), allowing the cumulative distribution function (CDF) to be estimated for each station. The CDF function was used to determine the non-exceedance probability of each timestep in the series, and, as the GEV was fit to annual timesteps, the annual exceedance probability follows a simple linear transformation with, for instance, a non-exceedance probability of 0.999, equivalent to a annual exceedance probability of 0.001. All values above this exceedance of 0.001 or a 1000-year return period were flagged as exceptionally high (Table 2 – Issue 3).

We note that real events will naturally fall above some of these thresholds. UK-Flow15 includes over 50,000 years of hydrometric data, and it is therefore expected that all these thresholds may be exceeded by genuine hydrological events.

Table 2: Mathematical representation for each high flow quality-control test used.

Nº	Quality control issue	Equation(s)
1	Above six standard deviations	$High\ flow\ 1: log(q_x + 0.01) > \overline{log(q + 0.01)} + 6SD[log(q + 0.01)]$
2	Double the second maximum flow	$Let\ AMAX_i = q_{\max\ at\ ranked\ year\ i}$ $High\ flow\ 2: \frac{AMAX_1}{AMAX_2} > 2$
3	Events above 0.001 annual occurrence	$CDF\ GEV_{\xi,\sigma,\mu}\ from\ \{AMAX_i\}$ $For\ each\ timestep: F_x = 1 - CDF\ GEV_{\xi,\sigma,\mu}(q_x)$ $High\ flow\ 3: F_x < 0.001$

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

Fitting stations to a statistical distribution and calculating the likelihood of a certain flow occurring at a station also allowed station intercomparison. This helps to determine whether a lower extreme event is plausible —for example, we can verify if a high flow is also observed at a nearby station or coincides with high rainfall.

The gauging stations in the UK are grouped into hydrometric areas, with each area representing either an integral river catchment with one outlet to the sea, or several contiguous similar river catchments with separate outlets (National River Flow Archive, 2014). These areas were used to group the stations' extreme events and to verify the occurrence of events at nearby stations. Events that exceeded a 0.1 annual likelihood were subjected to two checks within a three-day window: 1) check if a 0.5 annual likelihood event has happened in the region, within those days; 2) check if an event that was at least 20 times more likely to happen, happened within the region. If either condition was met, the event was validated; otherwise, it was flagged as failing the hydrometric area QC, as no event extreme enough, both in terms of ratio and intensity was found in the available nearby stations. Figure 4 illustrates this using an event from October 2002 within hydrometric area 09. A 0.08-probability event occurred at station 9001. Within three days, four other stations in the region recorded events with probabilities ranging from 0.10 to 0.25 - meeting condition 1; and with likelihoods ranging from three times less rare to similarly rare - meeting condition 2.

In a similar fashion, the values were also crosschecked with NRFA average catchment daily rainfall data for the catchment. Two checks were applied over the preceding three days: 1) check if any day of rainfall during the period exceeds Q99 before the occurrence of a '>0.1 flood event'; 2) check if a rainfall event at least 20 times more likely to happen has caused the rare extreme flood event. For instance, for the same October 2002 event happening at station 9001, on 22 Oct a total of 52.6mm of rain was recorded – a 0.14 event in the timeseries – satisfying condition 2. This value is also above the 99th rainfall quantile for the station (23.8mm), satisfying condition 1 (Figure 4).

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

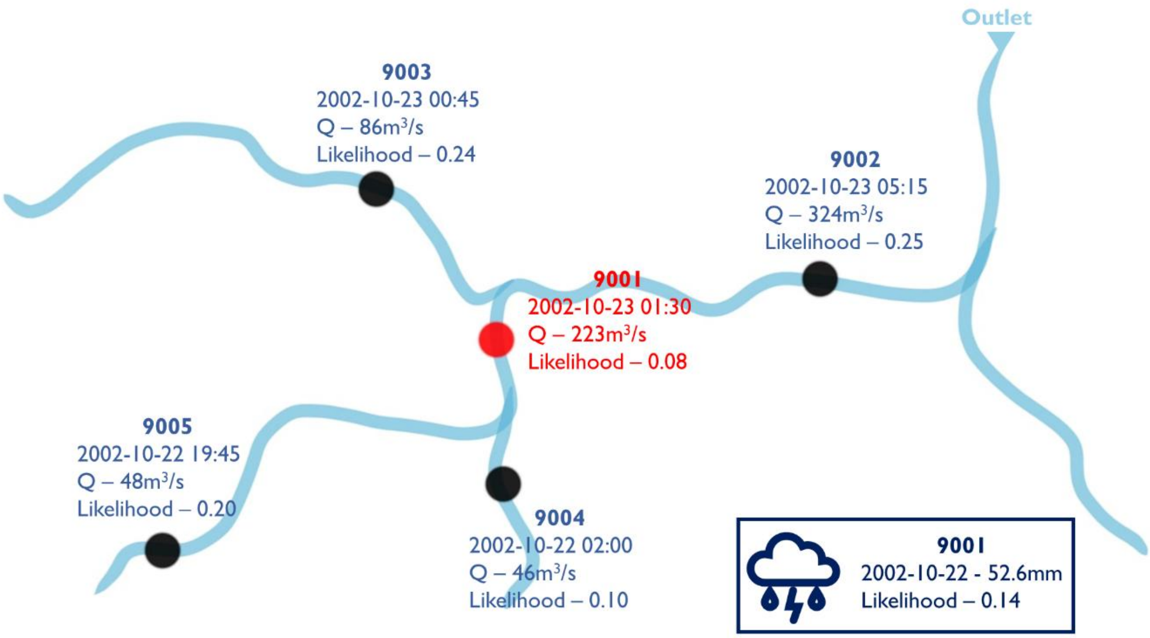


Figure 4: Example of the high flow quality-control by hydrometric region and by rainfall. Event at station 9001 - 23/10/2002 at 01:30.

4 Results

4.1 Visual inspection anomalies

Manual inspection identified 329 issues across 316 stations—approximately 23% of the dataset. Among these, 132 of the stations were flagged as having potential continuity issues – with 9 having issues in the entirety of the timeseries, 81 in part of the timeseries and 42 flagged as suspicious (Figure 5). Common causes included gauge changes (e.g., station 40027), vandalism (e.g., station 83009), and physical alterations such as reservoir construction (e.g., station 48009). Truncated stations correspond to around 3.5% of the dataset, often the consequence of rating limitation on the station. Some of the stations are known to have measurements that are imprecise at high values: this can be due either physical issues, such as bypassing of part of the flow once flows reach a limit (station 42001), or a limitation in flow calculations, due to the rating curve becoming imprecise at high flows (station 55003). 17 stations with fewer than five years of data were also flagged. For example, station 57014 still actively records data today but we have only received data from 2001 to 2004. Additionally, 91 stations contained isolated data clusters separated by large temporal gaps—for instance, station 22003 had 180 values in 1967, likely representing a flood event digitised from historical charts. These charts cover the period from 1957 to 1980 and may have been used in past flood analyses of the area. This was followed by a 32-year gap before a continuous record began in 1999 (Figure 5).

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

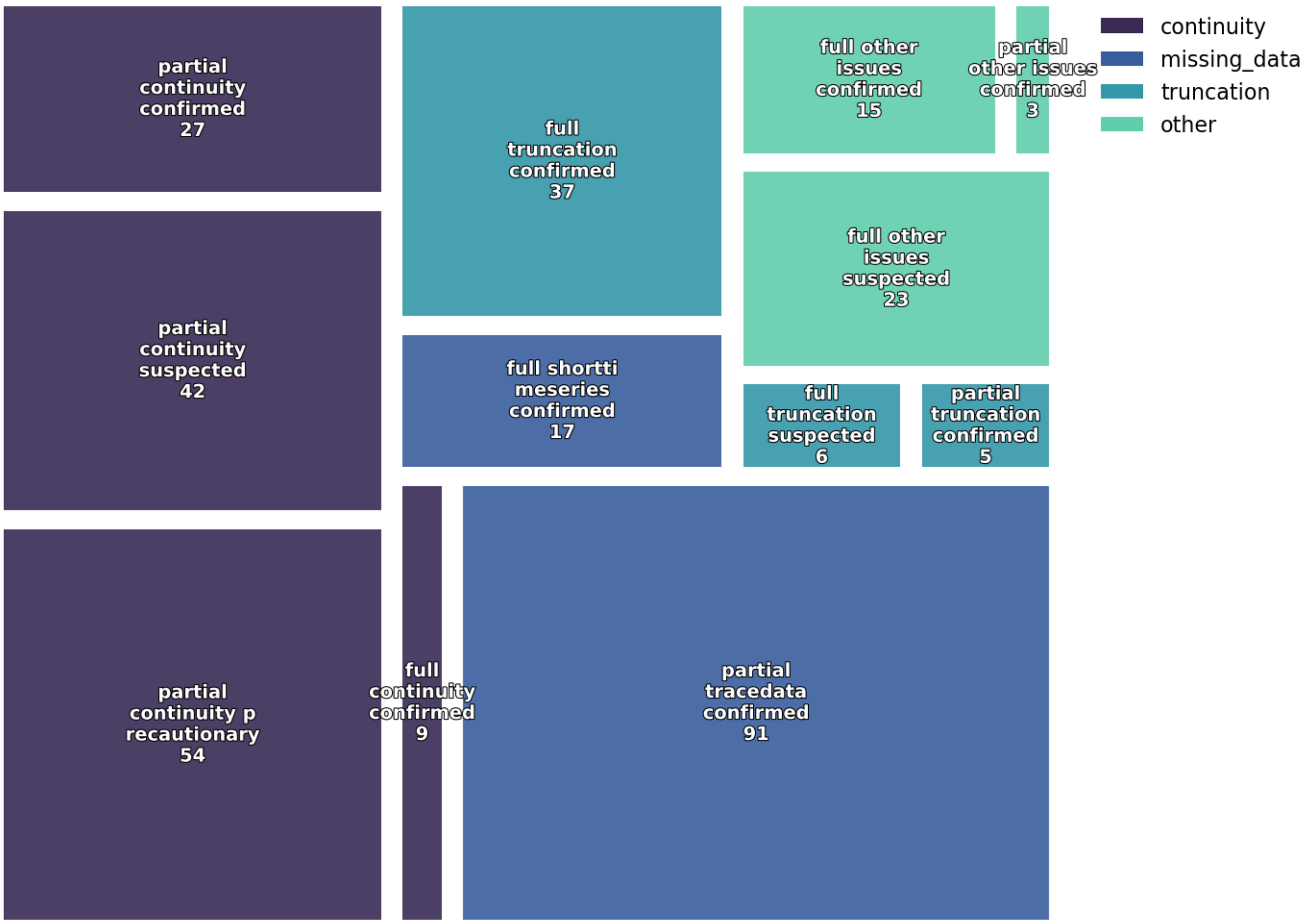


Figure 5: Visual inspection results - divided into four categories (colours) and types of errors (boxes).

4.2 Verification with other UK products

When compared with other UK datasets, the 15-min flow series generally shows good agreement with the NRFA daily dataset. However, 68 stations exhibited discrepancies for more than 10% of their values, and among these, 8 stations diverged from the NRFA for over 50% of their record. These inconsistencies can be attributed to several factors. One common cause is the use of combined stations within the NRFA daily dataset, where flows are aggregated from multiple gauging stations, an approach not reflected within the 15-min flow series (e.g., station 42010). Another source of discrepancy is different resampling conventions: in some cases, daily values have been computed over a 00:00 to 23:45 window, rather than using the standard hydrological window of 09:00 to 08:45 the following day (e.g., station 55034). The differences between NRFA AMAX values and the 15-min flow series are more notable. These typically result from corrections introduced during the manual review of annual maximum flow, or from older records where peak flows were assigned a default timestamp of 00:00, in contrast to the more precise temporal resolution available in the 15-min flow data.

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

A total of 66 stations contained duplicated timesteps with differing flow values. For 39 of these stations, the differences were minimal and occurred primarily during the transition period between irregularly timed records and standard 15-minute data, predominantly in 2003–2004. For these EA stations, the most recent data extracted from the API was retained, as it was the most recently published. The other 27 stations presented significant enough differences for individual checks, aiming to keep the most consistent value with other UK products, e.g. NRFA AMAX, POT and daily flows. The origins of these duplicates varied: in some instances, one record appeared time-lagged relative to the other; in others, discrepancies suggested digitisation errors, such as misplacement of decimal points. There were also cases where abrupt changes in flow (drops or rises) had been corrected in one version of the data but left unchanged in the other.

4.3 Traditional quality control

With the automated checks we have identified spurious data and unusual stations or behaviours. We envisage their utility as helping to provide a deeper understanding of the flow data. A total of 1070 stations have been flagged as having at least one flagged issue.

Negative values were present in 65 stations, most commonly with at least 96 occurrences - equivalent to one full day of data affected by artefacts (Figure 6). These are associated with ultrasonic stations or those influenced by tides. In ultrasonic stations, sudden drops may reflect sensor errors; in tidal stations, negative values may result from natural variations of the flow in the station.

Although only one station recorded a flow value exceeding 5000 m³/s, a step aiming to verify unrealistically high flows was deemed necessary. In total, 219 stations were flagged as having exceptionally high-flows, based on extreme thresholds, i.e., >6 standard deviations, an annual exceedance probability < 0.01, or double the second-highest annual maximum. In 75% of cases, these flow anomalies lasted less than a day. Prolonged flow anomalies often indicated serious issues, e.g., truncations or long periods of zero values. Spikes can also serve to identify unnaturally high flows and unnatural increases in flow. Absolute spikes were identified in 144 stations, and relative spikes in 508. Most affected stations had ≤4 flagged timesteps (Figure 6).

Drops and fluctuations in flow were very common occurrences in reservoir-influenced stations. In these stations, low flows sometimes present an oscillating series, followed by a drop/spike and subsequent oscillations around the new flow value. This behaviour leads to timeseries with long periods of fluctuations, as seen in Figure 6, where a large portion of the stations had fluctuations for more than a year. Drops are naturally less frequent, but still present in these stations, with the majority having less than a day of timesteps with drops (Figure 6).

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

Even though truncated low flows can be the consequence of faulty measurements, it is very common for ephemeral rivers to have these truncated measurements at either zero or very low values. Most low flow truncations lasted for a month, with some truncations lasting for more than 10 years (Figure 6). High flow truncations typically occur when the level gauge reaches its maximum capacity and begins to overspill, sometimes accompanied by flow bypass. We therefore recommend that truncated high-flow values be carefully verified before use in high-flow analyses.In the visual inspection anomaly checks we have flagged same value truncations along the timeseries, still, high truncations might occur at several thresholds in some stations.

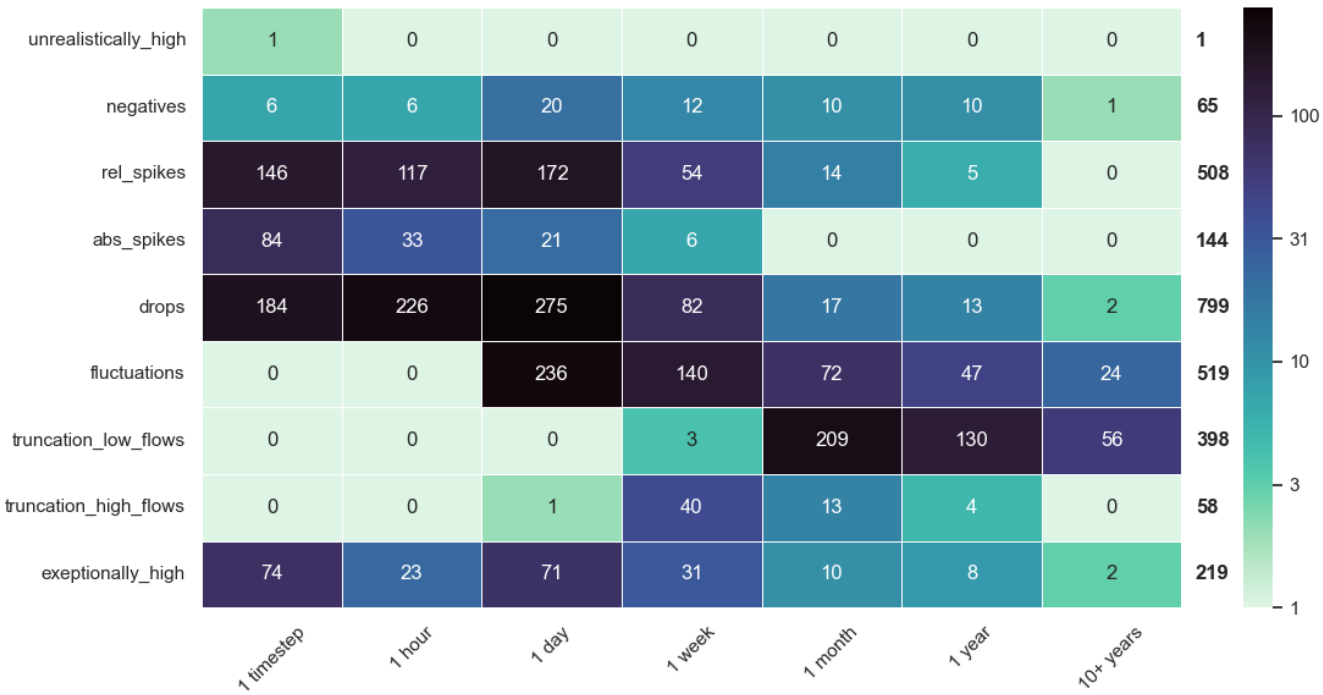


Figure 6: Summary of traditional QC flags - number of stations flagged per check (y axis) and per timesteps, converted into units of time (x axis). The colour scale representing the number of stations flagged on each category, with the total number of stations in the left corner.

4.4 High Flows Quality Control

The dataset contained 9,939 timesteps flagged as high flows with an estimated annual exceedance probability of 0.1—the threshold applied in both the hydrometric area and rainfall-based quality control procedures. This number is notably skewed by a subset of truncated stations, some of which had more than 1,000 flagged timesteps exceeding the 0.1 events/year threshold.

Approximately 10% of these flagged events were identified by the hydrometric area QC, indicating no concurrent high-flow event in any other station within the same hydrometric region. These cases were predominantly located in the south-east of

England (Figure 7a). In parallel, 14.5% of the flagged events were identified by the rainfall QC as lacking any preceding rainfall. These occurrences were more common in drier regions, while wetter areas of the UK—particularly the west—tended to show a stronger correspondence between rainfall and flooding, with many hydrometric areas exhibiting no events flagged as lacking prior rainfall (Figure 7b).

It is important to emphasise that such flags serve as useful indicators but are not, in isolation, definitive evidence of spurious events. The relationship between rainfall and flooding is complex and modulated by catchment conditions (Ledingham et al., 2019; Zheng et al., 2023), snowmelt, groundwater contributions, and other hydrological processes. Furthermore, genuine localised extreme rainfall events in small catchments may be incorrectly flagged as outliers by the regional QC framework. Events that were flagged by both criteria, i.e., lacking both preceding rainfall and any other event in the hydrometric region accounted for 3.4% of all events. These were disproportionately concentrated in the eastern part of the UK, particularly in chalk-dominated catchments (Figure 7c).

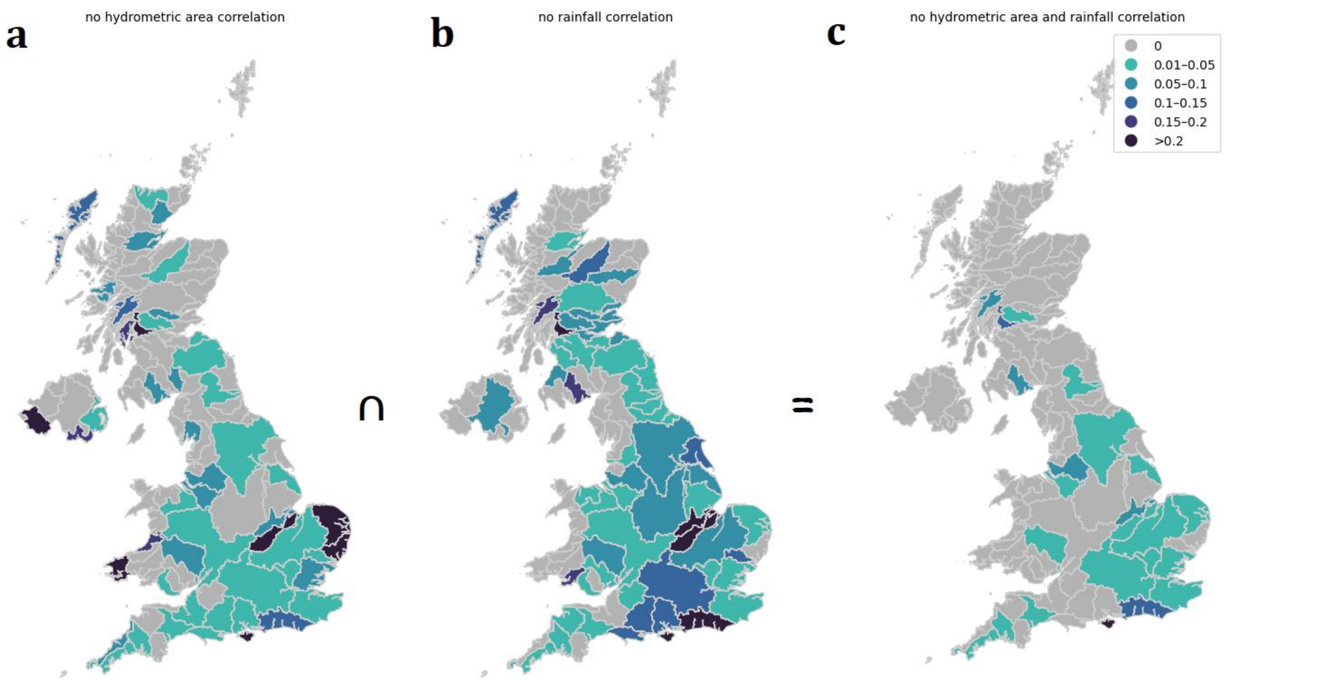


Figure 7: Outcome of the high flows QC. (a) Hydrometric areas that had intense events flagged solely in one station; (b) Hydrometric areas that had events that were not preceded by rainfall events; (c) hydrometric areas that had isolated events in one station that were not preceded by rainfall events.

5 Discussion

The availability of high-quality, sub-daily flow data is fundamental for hydrological modelling, as flow records are central to the calibration and validation of hydrological models (Beven and Binley, 1992; Knoben et al., 2019; Nash and Sutcliffe, 1970). Accurate model outcomes depend not only on model structure but also on the reliability of the input data (McMillan et al., 2018). In large-sample hydrology, where datasets may comprise hundreds or thousands of catchments, manually verifying all input time series for every scientific study is impractical. While the NRFA, aided by measuring authorities, maintains a dedicated team for manual quality control of the national peak and daily flow records, extending these checks to higher-resolution data would be exponentially more demanding and increasingly unfeasible. As a result, quality checks on high flows are often either rudimentary sometimes failing to detect important anomalies or done at a localized level failing to address the standardization needed in large-scale datasets.

Here, we present a national-scale, open-source QC framework tailored for the UK's 15-min river flow dataset, UK-Flow15 (Fileni et al., 2025b). The framework combines manual visual inspection of anomalies with automated detection of statistical artefacts, incorporating both established and novel procedures. In particular, we introduce methods to evaluate high-flow events by comparing them with rainfall records and flow observations from neighbouring catchments. Making a dataset available with pre-flagged quality concerns reduces the burden on users, facilitates the early identification of problematic inputs, and accelerates research workflows without compromising reliability. In addition, the framework could support the curation and maintenance of existing data archives and updates to other UK hydrological records.

5.1 Data quality impacts on hydrological modelling

In the UK, the NRFA has implemented a suite of QC procedures aiming on assuring the reliability of its daily flow products and allowing for accurate hydrological simulations at a national scale (Bell et al., 2009; Lane et al., 2019; Lees et al., 2021; Lewis et al., 2018b). Detailed metadata, such as records of gauge changes or reservoir construction, are often available for each station, however, some QC measures undertaken remain invisible to the user. Stations with poor data, (such as 84034 or 34020) are not available for download, despite having existing records. In other instances, data is partially excluded; for example, station 23016 is only available at daily resolution from the 1990s onward, despite having a longer record. Our approach to the 15-min dataset has been to retain the full timeseries and explicitly flag potential quality issues. This enables users to make informed decisions about data suitability based on the specific requirements of their analysis.

Even when rigorous procedures, such as those implemented in the NRFA daily series are applied, some data issues that may not have been deemed significant enough for removal can still affect model performance. We illustrate this using two large-sample hydrology studies that relied on daily NRFA data, which is itself derived from 15-min resolution measurements. For

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

consistency, we limited our comparison to stations where the NRFA daily data closely matched our 15-min data, allowing for no more than one year of discrepancy. Furthermore, we only considered stations with identified anomalies during the temporal window of each study. In Lane et al. (2019), six stations (23022, 27038, 28072, 31028, 33051, and 33062) had continuity issues flagged for the study period. These stations showed notable poor performance across the four lumped hydrological models evaluated, with a median Nash–Sutcliffe Efficiency (NSE) of 0.29 and a maximum NSE of 0.67, amongst all stations and models. Even machine learning models, such as LSTM and EALSTM, known for their efficient calibration and validation routine (Kratzert et al., 2018) are affected by poor quality data. In a UK national scale study (Lees et al., 2021), three (Stations: 41019, 41009, and 41004) of the five stations with the highest positive bias in high flows (Bias-FHV) were associated with a truncated time series, as identified through the visual inspection anomaly checks.

While these visual inspection checks remain effective for identifying major anomalies and discrepancies in hydrological data, automated procedures provide an additional layer of detail and consistency. The basic quality control (QC) framework developed enables the systematic detection and removal of unnatural values from the timeseries, as required by the user. Furthermore, the checks on high-flow events and comparisons with NRFA datasets offer a way of identifying suspicious events and cases of over- or underestimation. This is exemplified by station 58009 in the same study by Lees et al. (2021). During 1988 to 1994, most of the calibration period, the high flow values recorded were systematically lower than the corresponding NRFA POT and AMAX, indicating potential rating curve changes that have not been represented in the daily data. This underrepresentation of high-flow events during the calibration phase resulted in a pronounced positive bias (FHV) during the model's validation period.

It is important to note that the issues described above are observed in coarser, daily-resolution datasets, despite the smoothing effect introduced by temporal averaging. When working with sub-daily resolution data, the primary focus of this quality control procedure, these issues are expected to become even more pronounced. This further underscores the need for robust quality control methods capable of detecting and addressing such artefacts at finer temporal scales.

5.2 Limitations and future work

One of the primary challenges was defining appropriate thresholds for a national-scale dataset. Despite sensitivity testing, we acknowledge that both false positives and false negatives inevitably occur. For example, in the case of false positives, the rainfall QC will occasionally flag genuine events in groundwater-dominated catchments. Conversely, the timestep threshold for detecting high truncations often failed to identify problematic data, that were previously identified in the visual inspection anomaly check. To address these issues, threshold identification approaches tailored to catchment-specific characteristics, rather than one single national threshold, would be necessary.

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

The development of quality-controlled hydrological datasets is time and resource intensive. For instance, the GDSR sub-daily rainfall QC framework reflects over a decade of work (Blenkinsop et al., 2017; Lewis et al., 2018a, 2021), and the NRFA peak flow dataset is now in its 14th release, with annual reviews and updates (Turner et al., 2024). Nonetheless, in the context of flood risk and climate change assessment, minimizing uncertainty in input data remains essential. Errors in input data can propagate through the entire modelling chain: from hydrological simulation, where models introduce their own uncertainties (Gupta and Govindaraju, 2019; Lane et al., 2022; Renard et al., 2010; Teweldebrhan et al., 2018), to climate change projections and extreme flow extrapolations, which further amplify these uncertainties (Gao and Booij, 2020; Lane et al., 2022; Merz et al., 2022; Rodding Kjeldsen and Prosdocimi, 2023; Slater et al., 2021). Ensuring the reliability of hydrometric data as the initial driver of these analyses is crucial.

That is not to say there is no need to improve the quality of hydrometric data itself. In the UK we highlight two key actions: first, the extensive existing network of gauging stations must be maintained, with renewed efforts to enhance measurement precision through more measurements for stage-discharge validation and adoption of novel flow measurement techniques (Dolcetti et al., 2022; Manfreda et al., 2024); second, expanding the temporal extent of records by digitizing available historical datasets. The UK possesses tens of thousands of years of high-resolution historical flow data that remain undigitized. Digitisation efforts are essential and could build on successful precedents, such as the recovery of over three million rainfall observations dating back to the 17th century (Hawkins et al., 2023), and, on a smaller scale, the digitization of 150 years of high-resolution flow data in the Republic of Ireland (de Smeth et al., 2024).

6 Conclusion

We have produced a transparent and open-source quality-control framework for high-resolution (15-min) flow data. The QC framework integrates manual inspection of visual inspection anomalies with automated detection of statistical and hydrological artefacts. In addition to implementing established quality-control checks, we introduced new procedures to identify and classify anomalous high-flow events. The framework also addresses UK-specific data characteristics by assessing consistency with other national flow products. As a result, users can readily identify and interpret problematic or spurious data points.

Ensuring that hydrometric data are correctly managed, easily shared, and of good quality is a critical foundation for their reliable use in modelling and predictive analyses. An automated QC framework facilitates the identification, flagging, and, where necessary, exclusion of data that are unfit for specific analytical purposes. This is particularly important for historical observations, which were often originally produced solely to support real-time operational decisions. Implementing robust quality control is essential to realising the full value of the substantial investment made in generating these data.

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

While this framework marks a significant step forward in providing more reliable high-resolution UK flow data, we recognize that it is a foundation that could be further refined. Future research should focus on calibrating anomaly detection thresholds to reduce false positives and false negatives, by incorporating catchment-specific characteristics. Additional flags, such as comparisons with 15-min level data to verify rating curve conversions, to more complex checks such as assessment of at-station flood wave coherence, could further enhance quality assurance. Finally, a critical next step is the development of the quality-control code into a user-friendly package, which would support reproducibility, community contributions, and application to new datasets and regions. The availability of such tools is vital for modern hydrologyAs the processing of large datasets becomes increasingly feasible, reliable data have become crucial, especially under the pressures of climate change. Our framework directly addresses the urgent need for accessible and reliable high-resolution flow data, thereby supporting robust hydrological research, effective policy-making, and improved climate adaptation strategies in the UK and beyond.

Data availability

The dataset and metadata created in the first series of this papers are available at https://doi.org/10.5285/211710ac-f01b-4b52-807f-373babb1c368 (Fileni et al., 2025b)

Code availability

The quality-control code is available at https://github.com/felipef93/UK-Flow15-QC

Author contribution

FF: Conceptualization, Data Curation, Methodology, Formal Analysis, Investigation, Visualization, Software, Writing (original draft preparation); HJF: Conceptualization, Funding Acquisition, Methodology, Project Administration, Supervision, Writing (review and editing); EL: Conceptualization, Supervision, Methodology, Writing (review and editing); FM: Conceptualization, Supervision, Methodology, Writing (review and editing); GC: Supervision, Writing (review and editing); DA: Methodology, Supervision, Writing (review and editing); EB: Supervision, Writing (review and editing); LY: Conceptualization, Supervision, Methodology, Writing (review and editing); MF: Data curation, Supervision, Writing (review and editing); HC: Data curation, Writing (review and editing); OS: Data curation

Competing interests

The authors declare that they have no conflict of interest.

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

Acknowledgments

FF was part-funded by the Environment Agency's Flood Hydrology Improvements Programme and the Scottish Environment Protection Agency (SEPA) as a case partner for the project under the One Planet DTP (NE/S007512/1). HJF was supported by the Co-Centre for Climate + Biodiversity + Water, funded by UKRI (NE/Y006496/1). MF, HC, and OS were supported by the Hydro-JULES programme, funded by NERC (NE/S017380/1, NE/X019063/1). GC was supported by a United Kingdom Research and Innovation Future Leaders Fellowship (MR/V022857/1). This work was also supported by the Natural Environment Research Council (NERC-UKRI) as part of the Floods and Droughts Research Infrastructure (FDRI) project.

We clarify that neither the Environment Agency nor SEPA were involved in the development of the methodology or the review process. Still, the authors would like to express their sincere gratitude to the staff at the Department for Infrastructure – Northern Ireland, the Environment Agency, Natural Resources Wales, and the Scottish Environment Protection Agency for their valuable support in providing data and for their helpful responses to queries. We also extend our sincere thanks to the many hydrometrists, hydrologists, and others with expertise in UK flow data who shared valuable insights and engaged in interesting discussions with the authors.

References

Archer, D. and Fowler, H.: A historical flash flood chronology for Britain, J Flood Risk Manag, 14, https://doi.org/10.1111/jfr3.12721, 2021.

Di Baldassarre, G. and Montanari, A.: Hydrology and Earth System Sciences Uncertainty in river discharge observations: a quantitative analysis, Hydrol. Earth Syst. Sci, 913–921 pp., 2009.

Bauwe, A., Tiedemann, S., Kahle, P., and Lennartz, B.: Does the Temporal Resolution of Precipitation Input Influence the Simulated Hydrological Components Employing the SWAT Model?, J Am Water Resour Assoc, 53, 997–1007, https://doi.org/10.1111/1752-1688.12560, 2017.

Bell, V. A., Kay, A. L., Jones, R. G., Moore, R. J., and Reynard, N. S.: Use of soil data in a grid-based hydrological model to estimate spatial variation in changing flood risk across the UK, J Hydrol (Amst), 377, 335–350, https://doi.org/10.1016/j.jhydrol.2009.08.031, 2009.

Beven, K. and Binley, A.: The future of distributed models: Model calibration and uncertainty prediction, Hydrol Process, 6, 279–298, https://doi.org/10.1002/hyp.3360060305, 1992.

Blenkinsop, S., Lewis, E., Chan, S. C., and Fowler, H. J.: Quality-control of an hourly rainfall dataset and climatology of extremes for the UK, International Journal of Climatology, 37, 722–740, https://doi.org/10.1002/joc.4735, 2017.

Hydrométrie:

Campbell, A., Pradhanang, S. M., Kouhi Anbaran, S., Sargent, J., Palmer, Z., and Audette, M.: Assessing the impact of urbanization on flood risk and severity for the Pawtuxet watershed, Rhode Island, Lake Reserv Manag, 34, 74–87, https://doi.org/10.1080/10402381.2017.1390016, 2018.

Champion, M.: Les inondations en France depuis le VIe siècle jusqu'a nos jours: recherches et documents contenant les relations contemporaines, les actes administratifs, les pièces officielles, ets., de toutes les époques ..., V. Dalmont, 1861.

Chan, S. C., Kahana, R., Kendon, E. J., and Fowler, H. J.: Projected changes in extreme precipitation over Scotland and Northern England using a high-resolution regional climate model, Clim Dyn, 51, 3559–3577, https://doi.org/10.1007/s00382-018-4096-4, 2018.

Chan, S. C., Kendon, E. J., Fowler, H. J., Youngman, B. D., Dale, M., and Short, C.: New extreme rainfall projections for improved climate resilience of urban drainage systems, Clim Serv, 30, 100375, https://doi.org/10.1016/j.cliser.2023.100375, 2023.

Costabile, P., Costanzo, C., Kalogiros, J., and Bellos, V.: Toward Street-Level Nowcasting of Flash Floods Impacts Based on HPC Hydrodynamic Modeling at the Watershed Scale and High-Resolution Weather Radar Data, Water Resour Res, 59, https://doi.org/10.1029/2023WR034599, 2023.

Coxon, G., Freer, J., Westerberg, I. K., Wagener, T., Woods, R., and Smith, P. J.: A novel framework for discharge uncertainty quantification applied to 500 UK gauging stations, Water Resour Res, 51, 5531–5546, https://doi.org/10.1002/2014WR016532, 2015.

Crochemore, L., Isberg, K., Pimentel, R., Pineda, L., Hasan, A., and Arheimer, B.: Lessons learnt from checking the quality of openly accessible river flow data worldwide, Hydrological Sciences Journal, 65, 699–711, https://doi.org/10.1080/02626667.2019.1659509, 2020.

Dolcetti, G., Hortobágyi, B., Perks, M., Tait, S. J., and Dervilis, N.: Using Noncontact Measurement of Water Surface Dynamics to Estimate River Discharge, Water Resour Res, 58, https://doi.org/10.1029/2022WR032829, 2022.

Fileni, F., Fowler, H. J., Lewis, E., McLay, F., and Yang, L.: A quality-control framework for sub-daily flow and level data for hydrological modelling in Great Britain, Hydrology Research, 54, 1357–1367, https://doi.org/10.2166/nh.2023.045, 2023.

Fileni, F., Fowler, H. J., Lewis, E., McLay, F., Bruce, E., and Becker, M.: Flood Risk Assessments for Small Catchments Under Climate Change—How Can Scotland Improve Its Policy for Enhanced Flood Resilience and Preparedness, in: Journal of Flood Risk Management, https://doi.org/10.1111/jfr3.70035, 2025a.

Fileni, F., Fowler, H. J., Lewis, E., Fry, M., Cooper, H., Swain, O., Coxon, G., McLay, F., Bruce, E., Yang, L., and Archer, D.: Sub-hourly river flow data observations from 1369 river gauges in the UK, 1948-2023 (UK-Flow15), 2025b.

Gao, C. and Booij, M. J.: Assessment of extreme flows and uncertainty under climate change: Disentangling the uncertainty contribution of representative concentration pathways, global climate models and internal climate variability, Hydrol Earth Syst Sci, 24, 3251–3269, https://doi.org/10.5194/hess-24-3251-2020, 2020.

Green, A. C., Lewis, E., Tong, X., and Wardle, R.: A framework for incorporating rainfall data into a flooding digital twin, J Hydrol (Amst), 656, 132893, https://doi.org/10.1016/j.jhydrol.2025.132893, 2025.

Gudmundsson, L., Do, H. X., Leonard, M., and Westra, S.: The Global Streamflow Indices and Metadata Archive (GSIM)-Part 2: Quality control, time-series indices and homogeneity assessment, Earth Syst Sci Data, 10, 787–804, https://doi.org/10.5194/essd-10-787-2018, 2018.

Gupta, A. and Govindaraju, R. S.: Propagation of structural uncertainty in watershed hydrologic models, J Hydrol (Amst), 575, 66–81, https://doi.org/10.1016/j.jhydrol.2019.05.026, 2019.

Hawkins, E., Burt, S., McCarthy, M., Murphy, C., Ross, C., Baldock, M., Brazier, J., Hersee, G., Huntley, J., Meats, R., O'Grady, J., Scrimgeour, I., and Silk, T.: Millions of historical monthly rainfall observations taken in the UK and Ireland rescued by citizen scientists, Geosci Data J, 10, 246–261, https://doi.org/10.1002/gdj3.157, 2023.

Heuvelink, D., Berenguer, M., Brauer, C. C., and Uijlenhoet, R.: Hydrological application of radar rainfall nowcasting in the Netherlands, Environ Int, 136, https://doi.org/10.1016/j.envint.2019.105431, 2020.

Hou, J., Wang, N., Guo, K., Li, D., Jing, H., Wang, T., and Hinkelmann, R.: Effects of the temporal resolution of storm data on numerical simulations of urban flood inundation, J Hydrol (Amst), 589, https://doi.org/10.1016/j.jhydrol.2020.125100, 2020.

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

Huang, Y., Bárdossy, A., and Zhang, K.: Sensitivity of hydrological models to temporal and spatial resolutions of rainfall data, Hydrol Earth Syst Sci, 23, 2647–2663, https://doi.org/10.5194/hess-23-2647-2019, 2019.

Hulme, M. and Jenkins, G.: Climate Change Scenarios for the United Kingdom - UKCIP Technical Report No.1, 1998.

Kendon, E. J., Roberts, N. M., Fowler, H. J., Roberts, M. J., Chan, S. C., and Senior, C. A.: Heavier summer downpours with climate change revealed by weather forecast resolution model, Nat Clim Chang, 4, 570–576, https://doi.org/10.1038/nclimate2258, 2014.

Kendon, L., Short, C., Cotterill, D., Pirret, J., Chan, S., Pope, J., Jones, R., Lowe, J., Murphy, J., and Sobolowski, S.: UK Climate Projections: UKCP Local (2.2km) Transient Projections, 2023.

Kim, D., Chun, J. A., and Aikins, C. M.: An hourly-scale scenario-neutral flood risk assessment in a mesoscale catchment under climate change, Hydrol Process, 32, 3416–3430, https://doi.org/10.1002/hyp.13273, 2018.

Knoben, W. J. M., Freer, J. E., and Woods, R. A.: Technical note: Inherent benchmark or not? Comparing Nash–Sutcliffe and Kling–Gupta efficiency scores, Hydrol Earth Syst Sci, 23, 4323–4331, https://doi.org/10.5194/hess-23-4323-2019, 2019.

Kratzert, F., Klotz, D., Brenner, C., Schulz, K., and Herrnegger, M.: Rainfall–runoff modelling using Long Short-Term Memory (LSTM) networks, Hydrol Earth Syst Sci, 22, 6005–6022, https://doi.org/10.5194/hess-22-6005-2018, 2018.

Lane, R. A., Coxon, G., Freer, J. E., Wagener, T., Johnes, P. J., Bloomfield, J. P., Greene, S., Macleod, C. J. A., and Reaney, S. M.: Benchmarking the predictive capability of hydrological models for river flow and flood peak predictions across over 1000 catchments in Great Britain, Hydrol Earth Syst Sci, 23, 4011–4032, https://doi.org/10.5194/hess-23-4011-2019, 2019.

Lane, R. A., Coxon, G., Freer, J., Seibert, J., and Wagener, T.: A large-sample investigation into uncertain climate change impacts on high flows across Great Britain, Hydrol Earth Syst Sci, 26, 5535–5554, https://doi.org/10.5194/hess-26-5535-2022, 2022.

Ledingham, J., Archer, D., Lewis, E., Fowler, H., and Kilsby, C.: Contrasting seasonality of storm rainfall and flood runoff in the UK and some implications for rainfall-runoff methods of flood estimation, Hydrology Research, 50, 1309–1323, https://doi.org/10.2166/nh.2019.040, 2019.

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

Lees, T., Buechel, M., Anderson, B., Slater, L., Reece, S., Coxon, G., and Dadson, S. J.: Benchmarking data-driven rainfall-runoff models in Great Britain: A comparison of long short-term memory (LSTM)-based models with four lumped conceptual models, Hydrol Earth Syst Sci, 25, 5517–5534, https://doi.org/10.5194/hess-25-5517-2021, 2021.

Lewis, E., Quinn, N., Blenkinsop, S., Fowler, H. J., Freer, J., Tanguy, M., Hitt, O., Coxon, G., Bates, P., and Woods, R.: A rule based quality control method for hourly rainfall data and a 1 km resolution gridded hourly rainfall dataset for Great Britain: CEH-GEAR1hr, J Hydrol (Amst), 564, 930–943, https://doi.org/10.1016/j.jhydrol.2018.07.034, 2018a.

Lewis, E., Birkinshaw, S., Kilsby, C., and Fowler, H. J.: Development of a system for automated setup of a physically-based, spatially-distributed hydrological model for catchments in Great Britain, Environmental Modelling and Software, 108, 102–110, https://doi.org/10.1016/j.envsoft.2018.07.006, 2018b.

Lewis, E., Pritchard, D., Villalobos-Herrera, R., Blenkinsop, S., McClean, F., Guerreiro, S., Schneider, U., Becker, A., Finger, P., Meyer-Christoffer, A., Rustemeier, E., and Fowler, H. J.: Quality control of a global hourly rainfall dataset, Environmental Modelling & Software, 144, 105169, https://doi.org/10.1016/j.envsoft.2021.105169, 2021.

Lewis E., Quinn, N., Blenkinsop S, Fowler H J, Freer J, Tanguy M, Hitt O, Coxon G, Bates P, Woods R, Fry M, Chevuturi A, Swain O, and White S M: Gridded estimates of hourly areal rainfall for Great Britain 1990-2016 [CEH-GEAR1hr] v2, https://doi.org/10.5285/fc9423d6-3d54-467f-bb2b-fc7357a3941f, 2022.

Li, J., Li, L., Zhang, T., Xing, H., Shi, Y., Li, Z., Wang, C., and Liu, J.: Flood forecasting based on radar precipitation nowcasting using U-net and its improved models, J Hydrol (Amst), 632, https://doi.org/10.1016/j.jhydrol.2024.130871, 2024.
Manfreda, S., Miglino, D., Saddi, K. C., Jomaa, S., Eltner, A., Perks, M., Peña-Haro, S., Bogaard, T., van Emmerik, T. H. M., Mariani, S., Maddock, I., Tauro, F., Grimaldi, S., Zeng, Y., Gonçalves, G., Strelnikova, D., Bussettini, M., Marchetti, G., Lastoria, B., Su, Z., and Rode, M.: Advancing river monitoring using image-based techniques: challenges and opportunities, Hydrological Sciences Journal, 69, 657–677, https://doi.org/10.1080/02626667.2024.2333846, 2024.

McMahon, T. A., Peel, M. C., and Amirthanathan, G. E.: Assessing rating curve uncertainty, https://doi.org/10.1080/02626667.2024.2442736, 2025.

Mcmillan, H., Krueger, T., and Freer, J.: Benchmarking observational uncertainties for hydrology: Rainfall, river discharge and water quality, Hydrol Process, 26, 4078–4111, https://doi.org/10.1002/hyp.9384, 2012.

Mcmillan, H. K. and Westerberg, I. K.: Rating curve estimation under epistemic uncertainty, Hydrol Process, 29, 1873–1882, https://doi.org/10.1002/hyp.10419, 2015.

McMillan, H. K., Westerberg, I. K., and Krueger, T.: Hydrological data uncertainty and its implications, WIREs Water, 5, https://doi.org/10.1002/wat2.1319, 2018.

Merz, B., Basso, S., Fischer, S., Lun, D., Blöschl, G., Merz, R., Guse, B., Viglione, A., Vorogushyn, S., Macdonald, E., Wietzke, L., and Schumann, A.: Understanding Heavy Tails of Flood Peak Distributions, https://doi.org/10.1029/2021WR030506, 1 June 2022.

Nash, J. E. and Sutcliffe, J. V.: River flow forecasting through conceptual models part I — A discussion of principles, J Hydrol (Amst), 10, 282–290, https://doi.org/10.1016/0022-1694(70)90255-6, 1970.

National River Flow Archive: Integrated Hydrological Units of the United Kingdom: Hydrometric Areas with Coastline., 2014.

Renard, B., Kavetski, D., Kuczera, G., Thyer, M., and Franks, S. W.: Understanding predictive uncertainty in hydrologic modeling: The challenge of identifying input and structural errors, Water Resour Res, 46, https://doi.org/10.1029/2009WR008328, 2010.

Robson, A. J. and Reed, D. W.: Statistical procedures for flood frequency estimation. Volume 3 of the Flood Estimation Handbook., Wallingford, 1999.

Rodding Kjeldsen, T. and Prosdocimi, I.: Use of peak over threshold data for flood frequency estimation: An application at the UK national scale, J Hydrol (Amst), 626, https://doi.org/10.1016/j.jhydrol.2023.130235, 2023.

Schmidt, L., Schäfer, D., Geller, J., Lünenschloss, P., Palm, B., Rinke, K., Rebmann, C., Rode, M., and Bumberger, J.: System for automated Quality Control (SaQC) to enable traceable and reproducible data streams in environmental science, https://doi.org/10.5281/zenodo.8, 2023.

Slater, L., Villarini, G., Archfield, S., Faulkner, D., Lamb, R., Khouakhi, A., and Yin, J.: Global Changes in 20-Year, 50-Year, and 100-Year River Floods, Geophys Res Lett, 48, https://doi.org/10.1029/2020GL091824, 2021.

de Smeth, K., Comer, J., and Murphy, C.: Hydrometric data rescue and extension of river flow records: Method development and application to catchments modified by arterial drainage, Geosci Data J, 11, 176–196, https://doi.org/10.1002/gdj3.206, 2024.

Ladda ner observationer från sjöar och vattendrag:

Teweldebrhan, A. T., Burkhart, J. F., and Schuler, T. V.: Parameter uncertainty analysis for an operational hydrological model using residual-based and limits of acceptability approaches, Hydrol Earth Syst Sci, 22, 5021–5039, https://doi.org/10.5194/hess-22-5021-2018, 2018.

Tomkins, K. M.: Uncertainty in streamflow rating curves: Methods, controls and consequences, Hydrol Process, 28, 464–481, https://doi.org/10.1002/hyp.9567, 2014.

Turner, S., Sefton, C., Suman, G., Kumar, A., Swain, O., Barbedo, R., Smith, R., Griffin, A., Meresa, H., and Vesuviano, G.: National River Flow Archive NRFA PEAK FLOW DATASET VERSION 13 Note on changes from v12.1, 2024.

USGS Surface-Water Data for the Nation:

Wetterhall, F., He, Y., Cloke, H., and Pappenberger, F.: Effects of temporal resolution of input precipitation on the performance of hydrological forecasting, Advances in Geosciences, 29, 21–25, https://doi.org/10.5194/adgeo-29-21-2011, 2011.

Wilkinson, M. D., Dumontier, M., Aalbersberg, Ij. J., Appleton, G., Axton, M., Baak, A., Blomberg, N., Boiten, J. W., da Silva Santos, L. B., Bourne, P. E., Bouwman, J., Brookes, A. J., Clark, T., Crosas, M., Dillo, I., Dumon, O., Edmunds, S., Evelo, C. T., Finkers, R., Gonzalez-Beltran, A., Gray, A. J. G., Groth, P., Goble, C., Grethe, J. S., Heringa, J., t Hoen, P. A. C., Hooft, R., Kuhn, T., Kok, R., Kok, J., Lusher, S. J., Martone, M. E., Mons, A., Packer, A. L., Persson, B., Rocca-Serra, P., Roos, M., van Schaik, R., Sansone, S. A., Schultes, E., Sengstag, T., Slater, T., Strawn, G., Swertz, M. A., Thompson, M., Van Der Lei, J., Van Mulligen, E., Velterop, J., Waagmeester, A., Wittenburg, P., Wolstencroft, K., Zhao, J., and Mons, B.: Comment: The FAIR Guiding Principles for scientific data management and stewardship, Sci Data, 3, https://doi.org/10.1038/sdata.2016.18, 2016.

Yang, X., Liu, Q., He, Y., Luo, X., and Zhang, X.: Comparison of daily and sub-daily SWAT models for daily streamflow simulation in the Upper Huai River Basin of China, Stochastic Environmental Research and Risk Assessment, 30, 959–972, https://doi.org/10.1007/s00477-015-1099-0, 2016.

Yu, D., Xie, P., Dong, X., Hu, X., Liu, J., Li, Y., Peng, T., Ma, H., Wang, K., and Xu, S.: Improvement of the SWAT model for event-based flood simulation on a sub-daily timescale, Hydrol Earth Syst Sci, 22, 5001–5019, https://doi.org/10.5194/hess-22-5001-2018, 2018.

https://doi.org/10.5194/egusphere-2026-277
Preprint. Discussion started: 7 April 2026

Zheng, Y., Coxon, G., Woods, R., Li, J., and Feng, P.: Controls on the Spatial and Temporal Patterns of Rainfall-Runoff Event Characteristics—A Large Sample of Catchments Across Great Britain, Water Resour Res, 59, https://doi.org/10.1029/2022wr033226, 2023.

Zhou, Z., Smith, J. A., Baeck, M. L., Wright, D. B., Smith, B. K., and Liu, S.: The impact of the spatiotemporal structure of rainfall on flood frequency over a small urban watershed: An approach coupling stochastic storm transposition and hydrologic modeling, Hydrol Earth Syst Sci, 25, 4701–4717, https://doi.org/10.5194/hess-25-4701-2021, 2021.