BS 17898:2014

BSI Standards Publication

Code of practice for the management of observed hydrometric data

bsi.

...making excellence a habit.™

Publishing and copyright information

Published by BSI Standards Limited 2014

ISBN 978 0 580 81887 5

ICS 07.060

The following BSI references relate to the work on this document:
Committee reference CPI/113
Draft for comment 14/30277807 DC

Publication history

First published as BS 7898 August 1997
First published as BS 17898 December 2014

Amendments issued since publication

Date	Text affected

Contents

Summary of pages

This document comprises a front cover, an inside front cover, pages i to ii, pages 1 to 42, an inside back cover and a back cover.

Foreword

Publishing information

This British Standard is published by BSI Standards Limited, under licence from The British Standards Institution, and came into effect on 31 December 2014. It was prepared by Technical Committee CPI/113, *Hydrometry*. A list of organizations represented on this committee can be obtained on request to its secretary.

Supersession

This British Standard supersedes BS 7898:1997, which is withdrawn.

Use of this document

This British Standard takes the form of guidance and recommendations. It should not be quoted as if it were a specification and particular care should be taken to ensure that claims of compliance are not misleading.

Any user claiming compliance with this British Standard is expected to be able to justify any course of action that deviates from its recommendations.

Presentational conventions

The provisions of this standard are presented in roman (i.e. upright) type. Its recommendations are expressed in sentences in which the principal auxiliary verb is “should”.

Commentary, explanation and general informative material is presented in smaller italic type, and does not constitute a normative element.

The word “should” is used to express recommendations of this standard. The word “may” is used in the text to express permissibility, e.g. as an alternative to the primary recommendation of the clause. The word “can” is used to express possibility, e.g. a consequence of an action or an event.

Notes and commentaries are provided throughout the text of this standard. Notes give references and additional information that are important but do not form part of the recommendations. Commentaries give background information.

Contractual and legal considerations

This publication does not purport to include all the necessary provisions of a contract. Users are responsible for its correct application.

Compliance with a British Standard cannot confer immunity from legal obligations.

Introduction

Water management decisions and policies ought to be based upon quantitative knowledge of the hydrological system. Commonly, such knowledge results from observational hydrometric data, the collection of which is the subject of other standards, e.g. BS 7843 and BS EN ISO 18365. The subsequent management of such hydrometric data provides the linkage between field measurement and the eventual use of processed data to address a wide range of strategic and operational river and water management applications. As both the demand for and complexity of hydrometric data increase, it is important that the procedures and processes used to manage these data are standardized to allow greater intergration of data and ensure their protection for future use.

The availability of high-quality observational data is vital to developing an understanding of the hydrological cycle. Optimizing data management systems helps ensure that the maximum benefits are achieved from those resources invested in hydrometric monitoring. Effective standardized procedures for data transmission, manipulation, quality control, expression of uncertainty and storage are vitally important and their use is to be promoted throughout hydrometric observation networks.

Those responsible for hydrometric data management are encouraged by this code of practice to adopt the ethos of a professional stewardship and to remember their role as guardians of an important national, and sometimes international, resource.

This code of practice is concerned with general aspects of good practice in data management. Techniques for managing data are recommended, covering metadata collection, data storage and quality control. This standard assumes that the raw data have been collected and transmitted from the field in line with other British Standards for hydrometry, so this standard concentrates on the subsequent processing and management of these hydrometric data.

1 Scope

This British Standard gives recommendations for the management of observed hydrometric data, including raw data and other data and statistics derived from these observations. While the principles of data management can be applied to all hydrometric observations, particular focus is placed on measurements of precipitation, water level (including stage) and discharge in open channels.

NOTE The range of sites where water levels, and sometimes flow, are measured includes lakes, reservoirs, rivers, canals, tidal waters, sewers, wells, and boreholes.

There are other data (for example, evaporation, soil moisture and snow depths) that are embraced by the term "hydrometry", as defined in this British Standard, and there are yet others (for example, water temperature, pH, atmospheric pressure) that are peripheral. These are not discussed explicitly but it is intended that the same general principles be applied to the management of all such data.

The standard covers metadata associated with hydrometric data, including recommendations for the production and management of descriptive, analytical and statistical material relating to sites where hydrometric data are collected. The recommendations of this standard can be applied to some forms of data directly derived from observational records (for example, summary time series of monthly mean river flows), but are not designed for the management of data resulting from more complex numerical models or spatially aggregated datasets (for example, remotely-sensed data).

This standard does not cover the field collection of data or its transmission, but focuses on the management of data once they have been received in a hydrometric information management system.

2 Normative references

The following documents, in whole or in part, are normatively referenced in this document and are indispensable for its application. For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

BS 7843 (all parts), *Acquisition and management of meteorological precipitation data from a gauge network*

BS EN ISO 772, *Hydrometry – Vocabulary and symbols*

BS EN ISO 18365, *Hydrometry – Selection, establishment and operation of a gauging station*

BS ISO 1100-2, *Measurement of liquid flow in open channels – Part 2: Determination of the stage-discharge relation*

BS ISO 8601, *Data elements and interchange formats – Information interchange – Representation of dates and times*

3 Terms and definitions

For the purposes of this British Standard, the terms and definitions given in BS EN ISO 772 and the following apply.

3.1 data flag
indicator relating to the quality and characteristics of an observation

3.2 derived data
information calculated, or deduced, from raw data (**3.5**)

3.3 precipitation
water or ice derived from the atmosphere and deposited at ground level

NOTE Measured in terms of the depth in millimetres (mm) of its liquid equivalent.

3.4 quality control
process of confirming that the data held are a reliable representation of the variable being measured or derived

3.5 raw data
data resulting directly from the measurement of variables

3.6 UTC
Coordinated Universal Time

4 Principles of hydrometric data management

4.1 The requirement for data management

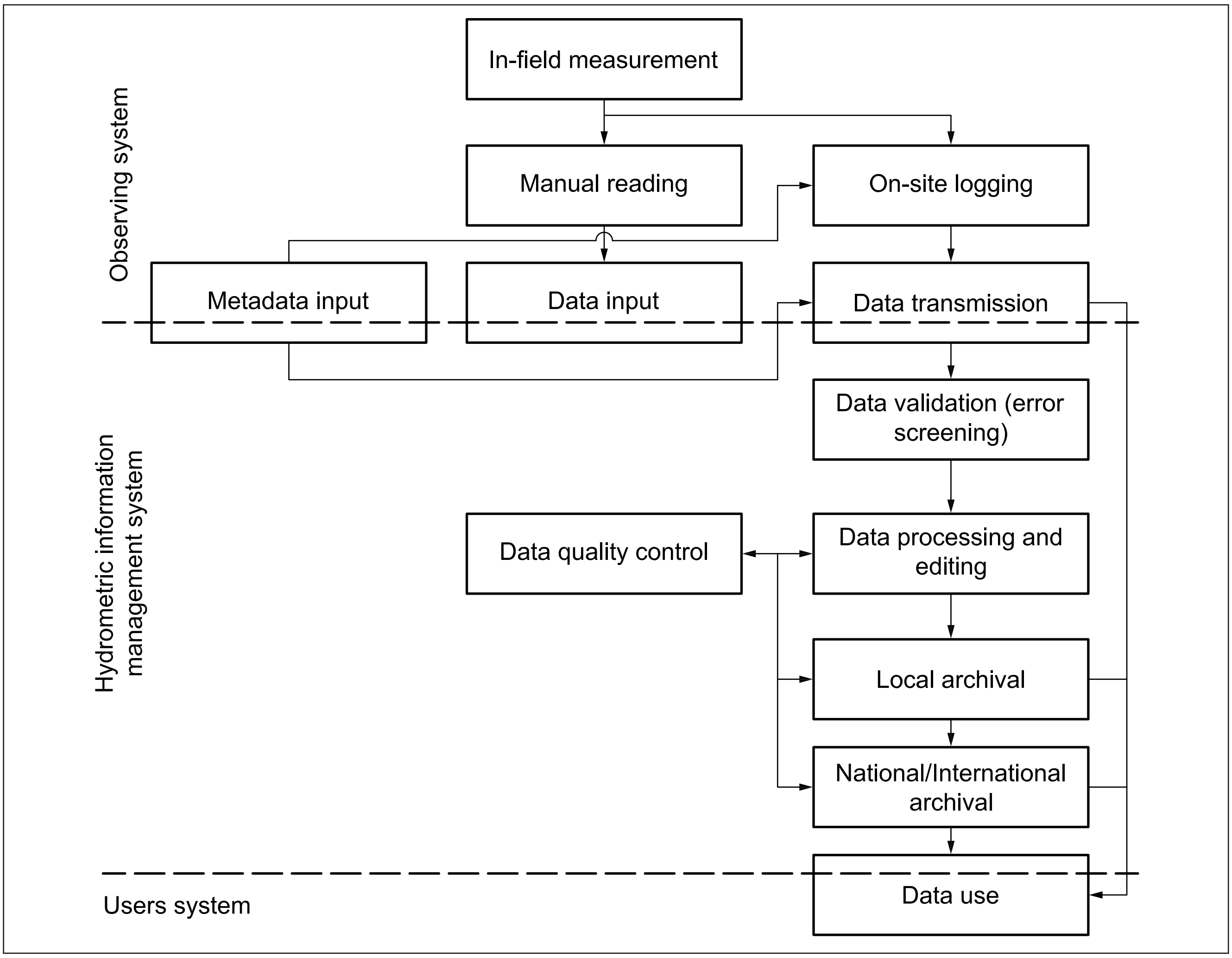
Hydrometric data management systems should be designed such that the utility of data to end-users at all levels is a central consideration. The value of data can be significantly affected by the treatment of data after collection, so data management systems should be designed to maximize the quality and continuity of records.

As water management decisions and policies are based upon quantitative knowledge of the hydrological system both now and in the past, the protection of data for future use should be a central requirement of any data management systems. As such, the design of such systems should consider, as far as possible, both current and future user requirements.

4.2 The sequence of data management

The management of hydrometric data is a continuous sequence (See Figure 1), which starts in the field at the point of collection and ends if and when data are destroyed. The design of any hydrometric data management system should consider all stages in this sequence.

Figure 1 **Data management sequence**



The management of data at the point of collection and their subsequent transmission vary depending on the hydrological variable being recorded, hydrometric methods employed and instrumentation used. Any data management systems used should ensure that adequate data and metadata are recorded at the point of collection to allow future management and use.

Some data management processes, such as quality control, may occur at more than one stage in the sequence and be completed by more than one organization. In designing, operating and reviewing data management systems, all stages in the acquisition, processing and dissemination of data should be considered.

4.3 Duty of care

4.3.1 General

Data often have significant value outside the organization that collected them, and data management systems should be designed, where possible, to capture all information required for future, sometimes unforeseen, uses.

Responsibility for the operation of hydrometric monitoring systems may be transferred (for example, to another organization or other personnel), so data management systems should be capable of preserving all data and associated metadata required to inform new operators.

4.3.2 Legacy data

The duty of care to manage hydrometric data should be taken to cover any such data currently held by an organization. In addition to data collected under the direct control of the organization, the duty of care extends to data collected by predecessors and inherited.

An assessment should be made of all legacy data upon inheritance (or discovery) to determine their contents and current state. An assessment should be made of their current, and likely future, utility to the organization and other external parties. While the management (and improvement) of legacy data can be demanding, the value of the resulting information can be high.

NOTE 1 Long historical records, even those of lower hydrometric quality, can be particularly valuable in providing an indication of previous hydrological conditions and often improve the value of current data through extending record lengths.

All legacy data should be actively managed. Where possible, such data should be brought into current data management systems and managed in accordance with **4.5** to **4.10** and Clause **5** to Clause **9**.

Where legacy data are not integrated into management systems for current data, they should be secured in archives maintained either by the organization which collected/inherited them or by independent bodies (for example, national archives, libraries).

Where the quality of legacy data (including inherited data) is poor, efforts should be made to improve it. Special attention should be paid to the collation and storage of metadata (both hard copy and digital) associated with data. In the case of discharge data, current meter gauging records and details of derived rating relationships are of particular value and should be maintained. Care should be taken to preserve operational catchment management information, for example, sluice gate operation, reservoir releases and resiting of raingauges, which is often vital for understanding apparent anomalies in hydrometric data.

Where data are considered "at risk" (for example, due to physical deterioration or storage in obsolete systems/formats) then consideration should be given to a programme of "data rescue" (see **4.8.7**).

Due to the difficulty in assessing future, currently unknown, requirements for hydrometric data, legacy data and associated information holdings should normally be preserved even if there is no immediate requirement to use the data.

*NOTE 2 Where data are deemed surplus to requirements, their archival (alongside metadata describing their lineage) is highly preferable to disposal. Recommendations for data disposal are given in **4.8.3**.*

4.4 Data types and terminology

For the purposes of **4.5** onward, hydrometric data should be considered to include:

a) time series of continuous or regular observations of a hydrological variable at a point-in-space;

b) sporadic or single point-in-time observations of a hydrological variable at a point-in-space;

c) metadata (including, discovery metadata and temporal metadata) and other descriptive material which provide information about the measurement, hydrological conditions and/or subsequent processing of hydrometric data.

Hydrometric data should also be taken to include time series and statistics derived from these data types. Typical examples of such hydrometric data include: total monthly rainfall, mean annual flood and Q_{95} (the flow equalled or exceeded 95% of the time).

NOTE 1 For the purposes of this standard data types a) and b), as well as time series subsequently derived from such data, are considered to be time series data.

NOTE 2 Spatially aggregated data (e.g. data produced by remote sensing or areal rainfall estimates) are not covered by this standard as their nature and management often differ from those used for other hydrometric data.

4.5 Time series data

4.5.1 Time of observation

All data points in a time series should be associated with a single point-in-time or period of time (e.g. for integrations), either by:

a) recording the start point-in-time and interval of a regular time series; or

b) recording the point-in-time against every value in a time series.

The representation of dates and times should conform to BS ISO 8601. When storing raw data, the standard Gregorian calendar conventions should apply, both to the year (1 January to 31 December) and day (midnight to midnight). To avoid gaps or overlaps in the records caused by biannual daylight-saving time adjustments, all times should be recorded in UTC. When recording an observation made at the instant dividing one calendar day from the next, then the time should be given as 00:00:00 (i.e. the start of the new calendar day).

Raw data should be recorded against the point-in-time to which the observation was made. Therefore, instantaneous observations should be recorded against the point-in-time the measurement was taken, and accumulations (e.g. rainfall totals) should be recorded against the end point of the accumulation.

Processed data (including derived data representing integration over a set period of time, e.g. daily mean flow or monthly rainfall total) may be recorded against either the start or end of the period in question. The convention used should always be clearly specified in the metadata associated with the record or a related data dictionary.

UK convention should normally be to calculate derived data using the hydrological day/month/year. Whether using these periods or others, the convention applied should be clearly specified in the metadata associated with the record or a related data dictionary. If using hydrological days/months/years (otherwise known as water or rainfall days/months/years) the following conventions should be used when deriving data:

1) hydrological day = observations made in the 24-hour period starting at 09:00 UTC, recorded against the start point of the period;

2) hydrological month = observations made during the hydrological days of that month; and

3) hydrological year = observations made during the hydrological days of 01 October to 30 September, inclusive, recorded against the end year of the period.

While the UK convention [as recommended in this subclause] is to present processed daily data (for example, total daily rainfall or daily mean flow) for the hydrological day, the raw data from which these time series are produced should always be stored using the normal calendar day. For example, raw hourly rainfall total data recorded at 06:00 on 13 April would be recorded against that time in the raw data series but subsequently included in the processed daily rainfall total for 09:00 on 12 April to 09:00 on 13 April, presented against the 12 April hydrological day. As a result, to avoid confusion, when presenting hydrological data for daily (or longer) time periods the associated metadata should always clearly specify whether they relate to calendar days/months/years or hydrological days/months/years.

4.5.2 Derived data terminology

Common examples of derived data produced from raw discharge time series are daily mean flows and annual maximum flows.

For any derived data or statistics held, the data management systems should clearly describe the method of derivation in metadata associated with the dataset. There should be clear, non-ambiguous associations between all derived datasets and the raw time series data from which they are derived.

While many derived data and statistics may be generated on demand, to assist in the future understanding of data processing, the data derived at each of the important stages in the data processing (or an unambiguous method for regenerating them) should be stored, even if these are only interim data. For example, when deriving daily mean discharge from a 15-minute stage series for an open-channel gauging station, it would be useful to store both the edited stage series, the derived 15-minute discharge and the subsequent daily time series.

In cases where statistical extremes are extracted at regular intervals prior to averaging (or vice versa), care should be taken to describe the resulting item of derived data correctly. For example:

a) maximum daily mean discharge (the maximum value of the sequence of daily mean discharges in a stated period);

b) mean daily maximum discharge (the average value, in a stated period, of a sequence of maximum instantaneous discharges recorded daily);

c) mean discharge (the average discharge over a stated period); and

d) mean annual runoff (the average value of annual flow volume expressed in millimetres over the catchment area).

4.5.3 Resolution of data storage

The number of significant figures to which data are stored should reflect the resolution of the original measurement (not necessarily the instrument), for example, staff gauge readings are usually read to 1 mm or 0.001 m. In the case of raw data, the storage precision should not degrade original measurement. The recommended maximum resolution for each data type are given in **6.3**, **7.1.5**, **8.3** and **9.4**.

4.5.4 Uncertainty

NOTE It is important that those responsible for hydrological data management systems have a good understanding of the component and overall hydrometric uncertainties pertaining to the data for which they are responsible.

ISO/IEC Guide 98-3, Guide to the expression of uncertainty in measurement (GUM), and DD CEN ISO/TS 25377, Hydrometric uncertainty guide (HUG), provide guidance on the estimation of uncertainties in hydrometric measurements and determinations.

A measurement result comprises:

a) an estimate of the measured value; and

b) a statement of the uncertainty of the measurement.

In accordance with DD CEN ISO/TS 25377 (HUG), the uncertainty of a hydrometric measurement or determination should be estimated as a combined uncertainty, calculated from the various component uncertainties. For example, the overall uncertainty in a flow determined from a stage-discharge rating should consist of the combined uncertainties in the stage measurement and the rating.

ISO/IEC Guide 98-3 (GUM) and DD CEN ISO/TS 25377 (HUG) use standard uncertainties (i.e. at the 68% confidence limit). DD CEN ISO/TS 25377 (HUG) recommends final resultant uncertainty of measurement to be expressed at the 95% confidence limit since this is normal hydrometric practice. Nevertheless, in the first instance, the standard uncertainty is estimated and this result is multiplied by the coverage factor ($k = 2$) to express the uncertainty at the 95% confidence limit. The statement of the uncertainty should specify the confidence level or number of standard uncertainties for which the uncertainty estimation has been made.

4.6 Metadata and descriptive material

The metadata and other descriptive material which provide information about the measurement and/or subsequent processing of hydrometric data are wide ranging and can take many forms. Clause **5** provides further recommendations relating to metadata, but the following principles should be observed.

Metadata should be stored permanently and organized in such a way that they can be linked to any measurement to which they relate. For example, where the information being stored relates to a specific monitoring site this should be associated with the site's unique identification code in a defined referencing system (see **5.2.2**).

Where information relates to a specific time (or time period) this should be clearly recorded alongside the information. For example, the period of applicability of a particular stage-discharge relation should be stored along with the definition of the relation.

Each piece of information should have a record of the name/affiliation of the author and date of preparation. The same information should be recorded for any subsequent changes to information.

Processes should be in place to ensure that metadata records are updated.

4.7 Maximizing data utility

NOTE The recommendations of this subclause are intended to ensure that hydrometric data management systems are designed and operated to maximize the utility of data both now and in the future.

4.7.1 Availability of metadata

All hydrometric data should be associated with descriptive metadata detailing their origin and subsequent processing. Such information should be made available to data users.

4.7.2 Data quality control

Hydrometric data should be subject to defined quality control procedures to ensure that they are fit for purpose. Quality control procedures should cover the assessment and subsequent alteration of data. Quality control normally involves editing and applying quality flags and comments to improve the quality and utility of the data, and should be undertaken to:

a) minimize instrument errors;

b) minimize processing errors and human errors; and

c) ensure the data and their derivatives are plausible.

Any changes to data should be made in accordance with **4.8.2**.

NOTE 1 Where more than one organization is involved in processing data, it is essential that responsibilities for the quality control and changing of data are clearly agreed by the organizations concerned.

Quality control procedures for hydrometric data should be defined according to the following four levels, with the choice of quality control level ranging from 1 (no quality control) to 4 (extended quality control) reflecting the importance of the data to anticipated users.

- Level 1: No quality control.
- Level 2: Primary quality control.
- Level 3: Secondary quality control.
- Level 4: Extended quality control.

NOTE 2 Quality control procedures may incorporate more than one level of checks according to a predefined schedule. For example, a site may receive Level 2 checks every week, Level 3 checks every month and Level 4 checks every year.

*NOTE 3 The techniques for quality control of hydrometric data vary according to the parameter being observed and further recommendations for what each level of quality control should involve are given in Clause **6** to Clause **9**.*

4.7.3 Data gap infilling

Gaps in time series data can severely damage the utility of records and should be minimized. Where appropriate, estimates of missing data should be made and these estimates included in the processed dataset. Data infilling should be done at the appropriate time step.

*NOTE Guidance on methods for data infilling can be found in **7.2.4** and **8.5.3**.*

Estimated data should always be clearly flagged as such (see **5.5.3**) and should be associated with metadata describing how the estimate was made, by whom and when.

4.7.4 Regular data reviews

Regular reviews of all data currently collected or stored (including legacy data) should be carried out to assess the degree to which the data management system and procedures are meeting the identified user requirements and to identify any necessary improvements.

Regular reviews of individual time series should be carried out to identify and address issues such as: recurring problems encountered during validation, associating apparent changes in data with physical changes at a site/in the observation methodology, the implications of site maintenance, and ensuring metadata are complete and up-to-date.

4.8 Data storage considerations

4.8.1 Data storage

In order that reviews and investigations into past hydrological events can be properly conducted and allow a full understanding of historical monitoring, all raw observed data should be permanently stored. These may be in hard copy format or digital form, depending on the original observation method (see **4.8.5**).

Erroneous data should be stored but flagged as such to prevent its future misuse. However, any data judged to be entirely unrepresentative of the measured variable (for example, as a result of known severe system problems) may be discarded.

Data and statistics derived from raw data may be stored or computed on demand in accordance with **4.5.2**.

4.8.2 Changes to data

In accordance with **4.8.1**, raw data should never be destroyed or amended. To preserve raw data, any changes or corrections to original data should be made within a parallel data series and should be flagged, recorded and described in metadata associated to the altered data point(s). If changes are applied to a whole record (for example, datum shift) this should be recorded in the metadata for the dataset.

Where changes are made to any data, a record should be kept of the change along with details of who it was made by and when.

4.8.3 Data disposal

Data archival (alongside metadata describing their lineage) is highly preferable to disposal. Where data disposal is proposed or data are at risk, other potentially interested parties should be identified and contacted to determine whether data could be transferred elsewhere.

If any data are disposed of, records of what was destroyed, when and why should be kept.

4.8.4 Electronic data storage system considerations

Electronic data can be stored in a range of different digital storage systems, from flat files to dedicated relational databases. When choosing a data storage system consideration should be given to security, ease of maintenance, costs, accessibility, possible future expansion, and ability to store metadata and other associated data/information.

Where possible, data storage systems should not be reliant on particular software. The data should be held in a format which can be easily accessed and transferred into a new system if software being used to manage it becomes obsolete or unavailable.

All data storage systems used should have the ability to contain both the hydrometric data and the associated metadata (including data flags).

A catalogue of the current contents of any database or file system should be maintained.

Where possible, a record of changes to the data storage system should be maintained detailing what changes were made, by whom and when.

4.8.5 Tapes, charts and other data in paper format

Some raw data, especially historical records, can exist in the form of paper charts or tapes. Such data should be digitized to maximize their utility before they become unreadable due to physical deterioration or inaccessible due to lack of equipment to read the data format. Where the original format contains key information not captured during the digitizing/scanning process, the originals should be stored in a secure and catalogued fashion.

4.8.6 Data security

Data management systems should include backup and data recovery systems appropriate to the importance of the data being stored.

Processes should be put in place to ensure that the risk of the loss or corruption of data held in the data storage system is kept to an absolute minimum. Actions that should be taken include:

a) ensuring that the data management system and the computer system on which it resides meet up-to-date security standards;

b) ensuring the computer system is designed in such a way as to minimize the risk of infection by viruses and other malicious software;

c) ensuring those given write permissions to the data storage system are fully aware of the risks of corruption and the number of people with write access is kept to a minimum consistent with efficient operations;

d) undertaking regular monitoring of the pattern of changes to the data held in the system;

e) making regular backups of the system;

f) maintaining off-site backups, which may include archives for local use;

g) regularly reviewing the effectiveness of security procedures; and

h) where data are stored using systems or media which are at risk of becoming obsolete, transferring to a more secure storage system should be prioritized (see **4.8.7**).

4.8.7 Data rescue

A process of data rescue to improve data storage conditions (see **4.8.5**) should be considered where data are at risk of being lost due to deterioration of records (for example, by damage to the medium), or there is a high risk of catastrophic loss (due, for example, to storage media becoming unreadable, obsolete electronic file formats or loss of knowledge about records leading to unintentional destruction). A full inventory of the data holding (including measured values and metadata) should be made to identify the contents and current state of the data. Where data at risk of loss are not to be rescued the recommendations of **4.8.3** should be followed.

4.9 Archiving hydrometric data: national collations

NOTE 1 For the purposes of this standard, the archival of data refers to their long-term storage for future use.

Permanent national archives of precipitation, river discharge and groundwater level data for the UK should be maintained. Some network operators may maintain a regional or local archive which contributes data to the national archives.

The national archives should:

a) include current and historical measurements from as many appropriate sites as possible;

b) provide a complete set of associated metadata;

c) provide assurance of the quality of the measurements, including data flags; and

d) provide ease of access to the data.

A procedure should be undertaken at regular intervals to determine which monitoring sites contribute to the national archives. The national archives should contain data relating to all sites that make, or have in the past made, regular measurements of the variable. Sites may be excluded from the archives if:

1) the data are considered to be of unacceptable quality (e.g. due to poor observing standards); or

2) the owner of the intellectual property does not allow general use of the data.

All operators of precipitation, river discharge and groundwater level monitoring sites should consider contributing the resulting data to the national archives.

The archive of current data should be in electronic format, with additional paper records where full digitization and quality control have not been completed.

To ensure the national archives remain up to date, data should be transferred as soon as they have been validated by the collecting organization and then updated when data are changed. Appropriate metadata should be transferred and stored alongside time series.

Efforts should be made to ensure appropriate data consistency between national archives and regional/local archives where they exist.

NOTE 2 The data in the national rainfall archive are designated as permanent, as defined by the Public Records Act 1958 [1]. The Act specifies approved places of deposit for historical paper records.

NOTE 3 All precipitation data in the national archive since 1961 are in electronic format. Many older records, some dating back to the 19th century, are only available in manuscript format. However, a complete electronic record exists for selected long-period stations.

4.10 Electronic transfer of data

Electronic hydrometric data, including metadata (see **5.2.1**), should be transferred in a text-based format. The key metadata outlined in Clause **5** should always be included in the data format.

Where possible, transfer of data in proprietary formats should be avoided. While other data formats may be used when transferring data from one machine/data system to another, preference should be given to exchanging hydrological data in an electronic data format currently specified by an ISO, World Meteorological Organization (WMO) or Open Geospatial Consortium (OGC) standard.

The most appropriate internationally agreed format is for data to be made available for transfer in the latest OGC-approved XML format for in situ hydrological time series. Currently, this is the *WaterML2.0 Part 1: Time Series Encoding Standard* produced jointly by OGC and WMO [2]. Related WaterML standards should be used for other types of hydrological data where these are available (for example, using the ratings and gauging standard currently under development).

The format for international exchange of precipitation data should conform to WMO [3 and 4] recommendations.

NOTE The principal formats for the international exchange of precipitation data are SYNOP (synoptic surface observations), SREW (synoptic rainfall Europe west), BUFR (binary universal form for the representation of meteorological data) and CREX (character form for the representation and exchange of data).

Where old data transfer systems operate with obsolete data formats a change should be implemented at the earliest convenient opportunity.

5 Metadata

5.1 General

Metadata and other descriptive material which provide information about the measurement and/or subsequent processing of data are vital to the understanding of hydrometric data. Such data should be actively managed in accordance with **4.3**.

The recommendations of this subclause for the generic metadata requirements should be applied to all hydrometric data. Where additional metadata are required for specific types of hydrometric data, these should be managed in accordance with Clause **6** to Clause **9**.

For the purposes of this standard, metadata (and other descriptive material) should be categorized into four groups.

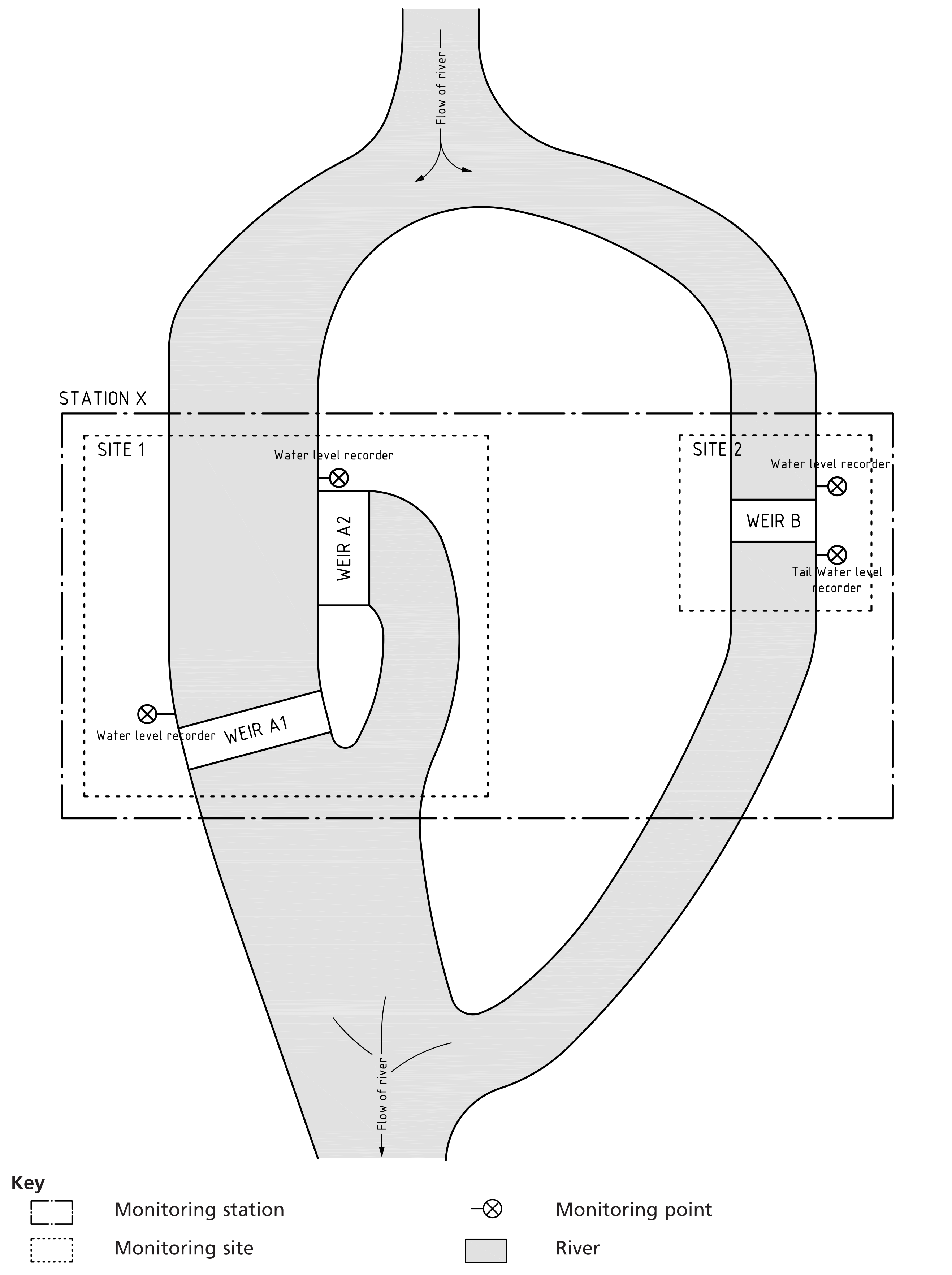
a) Monitoring station metadata: Information relating to a site or group of sites which are used to assess one or more related variable. For example, information about two sites operated in tandem to measure flow in different channels of the same river (see Figure 2).

b) Monitoring site metadata: Information relating to a set of collocated monitoring points which form a coherent single monitoring site. For example, information about a river discharge gauging station which comprises a number of weirs (see Figure 2) or one meteorological station containing multiple raingauges.

c) Monitoring point metadata: Information which relates to a specific measurement point (or instrument) located within a monitoring site/station. For example, information about a particular stage measurement taken at a river discharge gauging station.

d) Time series/Observation metadata: Information which relates to a measurement at a specific time (or over a particular time period). For example, information identifying an individual measurement as being erroneous.

Figure 2 **Example of monitoring stations, sites and points at a river discharge gauging station**



5.2 Monitoring station metadata

5.2.1 General

Metadata relating to the station at which the observation was recorded represent some of the most important information required when using hydrological data and should be managed with the same level of importance as the time series to which they relate.

There should be a clear and unambiguous link from each dataset through to the monitoring station to which it relates.

The metadata required to understand and interpret associated hydrological data varies dependent on the type of data and use. However, the following metadata should be recorded for all hydrometric monitoring stations:

a) unique station identifier(s) (see **5.2.2.2** and **5.2.2.3**);

b) UK hydrometric area (see **5.2.3**);

c) station name and water body name (where applicable) (see **5.2.4**);

d) station description (see **5.2.5**); and

e) type of station (see **5.2.6**).

5.2.2 Unique identifier

5.2.2.1 General

Each monitoring station/site should be given a unique identifying number (or code), which should be used to denote all data and other information pertinent to monitoring at that location.

5.2.2.2 Numbering systems

Each monitoring station/site within a network should be designated a unique code/number in a defined referencing system.

Where a station has a unique reference code/number in more than one network, the network system to which the identifier relates should be clearly stated.

A catalogue of all unique identifiers within any network should be maintained. Once an individual reference code/number has been designated to a site, this identifier should not be reused in future for a different station/site, even if the initial station/site closes.

5.2.2.3 UK national numbering systems

National numbering systems exist within the UK for river discharge gauging stations, raingauges and groundwater monitoring boreholes, and only stations/sites registered in the national networks should have a designated national number. The unique station/site identifiers within these systems should only be assigned by the body responsible for maintaining the national database in question.

NOTE UK national gauging station numbers are unique six-digit reference numbers (e.g. 039016) which serve as the primary identifier for the station record on the National River Flow Archive. The first three digits identify the hydrometric area in which the station is located (see ***5.2.5****), followed by a unique three-digit code. Leading zeros are normally not shown.*

UK national raingauge numbers are unique six-digit reference numbers, referred to as the Rainfall Site Number. The first digit is a regional identifier based on numbering from 0 to 9 in a clockwise direction around the British coastline from north east England. Northern Ireland and Channel Islands are included in region 9. The second and third digits identify the river catchment in which the gauge is located.

UK national groundwater well numbers are based upon the National Grid. Each 100 kilometre square is designated by prefix characters, e.g. SE, and is divided into 100 squares of 10 kilometre sides designated by numbers 00 (in the south-west corner) to 99 (in the north-east corner). Thus, the site SE94/5 is the fifth well or borehole recorded in the National Well Record Archive within the 10 kilometre square SE94. A suffix such as A, B, C defines the particular well when there are several at the same site. For Northern Ireland, where the Irish Grid is used, the first of the prefix characters is always "I".

5.2.3 UK hydrometric areas

The Hydrometric Areas for Great Britain and Northern Ireland are administrative groupings of river catchments at a sub-national level and, where appropriate, should be used in the organization of UK river flow records and hydrometric data management.

NOTE In the UK, hydrometric areas are defined as either integral river catchments having one or more outlets to the sea or tidal estuary or, for convenience, include several contiguous river catchments having topographical similarity with separate tidal outlets.

The hydrometric area identifier is a three-digit number. The first digit is a regional identifier being 0 for mainland Britain, 1 for the larger islands/groups of islands around Britain and 2 for Ireland.

The next two digits identify the hydrometric area within these regions. In mainland Britain the hydrometric areas are numbered from 1 to 97 in clockwise order around the coast, starting in north-east Scotland. Ireland has a unified numbering system for hydrometric areas but not all have an outlet to the coast.

The definitive number, name and geographical extent of each UK hydrometric area should be maintained by a single organization to avoid confusion.

5.2.4 Station and water body naming

The name of the measurement station (and water body, where applicable) should be standardized and recorded in the data management system in sufficient detail to avoid ambiguity. Where standard catalogues of water bodies exist these should be used.

5.2.5 Station description

A summarized general description of station which contains details of the location, hydrological characteristics, operational procedures, hydrometric limitations, characteristics of the local geographical area, and general comments on the record should be maintained. This should include changes over the life of the station so as to provide a historical description of the region it represents.

5.2.6 Type of station

The type(s) of monitoring conducted at the station should be recorded. The monitoring station type(s) should be selected from the following:

a) precipitation;

b) surface water; and

c) groundwater.

Where a station comprises multiple types, or monitors more than these, this should be recorded in the metadata.

5.3 Monitoring site metadata

5.3.1 General

NOTE Metadata relating to the individual site(s) at which a measurement was taken are important to aid understanding of the source of data.

As a monitoring station may contain one or more site, there should be an unambiguous link between the station identifier and that of the monitoring sites that form part of it. Where sites are combined into one station, the relationship between the sites should be detailed in the metadata. For example, where discharges recorded at two sites are totalled to give the overall river flow for a station, the method for doing so should be recorded in the metadata.

The metadata required to understand and interpret associated hydrological data varies dependent on the type of data and use. However, the following metadata should be recorded for all hydrometric monitoring sites:

a) unique site identifier(s) (see **5.2.2**);

b) geographical location (see **5.3.2**);

c) operating period (see **5.3.3**);

d) site description (see **5.3.4**);

e) site type (see **5.3.5**);

f) operator details (see **5.3.6**); and

g) additional site information (see **5.3.7**).

5.3.2 Geographical location

Within the site metadata, the geographical location should relate to the main measurement point within the site, the site benchmark or another defined location. The physical point to which the horizontal location and altitude relate (e.g. gauge zero, weir crest, top of raingauge) should be specified in the metadata.

The horizontal location, accurate to within 1 m, should be recorded using the Ordnance Survey national/Irish grid reference, or latitude and longitude.

The altitude of the monitoring point, accurate to within 1 m, should be recorded as the height above mean sea level (to the appropriate national ordnance datum).

The coordinate reference system and altitude datum used should be recorded in the metadata associated with the site. Efforts should be made to ensure consistent practice across a network and that this is recorded.

5.3.3 Operating period

The opening and, if applicable, closing dates of the site should be recorded. These should normally reflect the first date on which the monitoring at the site was considered to be operational and the date on which monitoring ceased.

5.3.4 Site description

Where a station contains more than one site a description of each component should be recorded. The description should include explanation or the relationships between the difference sites.

5.3.5 Type of site

The type of monitoring site should be recorded. Suggested classifications for monitoring site types are given in Clause **6** to Clause **9**.

5.3.6 Operator details

Details of the current site operator/owner/reporting authority should be maintained.

NOTE Attention is drawn to the requirements of the Data Protection Act [5] when handling personal details.

5.3.7 Addtional site information

Additional information about the site and its history can be of assistance to users of the associated hydrometric data. In particular, photographs, maps/plans and diagrams to illustrate the station’s location, set-up, operation or performance can greatly aid understanding of monitoring practice. The source of information should be recorded. The geographical location and time at which any illustrative material was produced (for example, the date a photograph was taken) should be provided.

5.4 Monitoring point metadata

5.4.1 General

Metadata relating to the individual monitoring point/instrument used to obtain a measurement is important to aid understanding of the resulting data.

Each site may contain one or more monitoring points/instruments. There should be an unambiguous link between the site identifier and that of the monitoring points/instruments that are part of it. Where multiple monitoring points/instruments are present on any one site, the relationship between instruments and site should be detailed in the metadata. For example, a river flow monitoring site may contain a primary water level monitoring point upstream of a control and a secondary tail water level monitoring point used for the correction of calculated discharges during drowned flow conditions.

The primary monitoring point/instrument within a site for each parameter at any one time should be clearly identified.

The metadata required to understand and interpret associated hydrological data varies dependent on the type of data and use. However, the following metadata should be recorded for all hydrometric monitoring point/instruments:

a) unique monitoring point/instrument identifier(s) (see **5.4.2**);

b) geographical location within the site (see **5.4.3**);

c) operating period (see **5.4.4**);

d) monitoring point/equipment details (see **5.4.5**); and

e) operational details (see **5.4.6**).

5.4.2 Monitoring point/instrument identifier

Each monitoring point/instruments within a site should be designated a unique code/number in a defined referencing system (see **5.2.2**).

A catalogue of all monitoring point/instrument identifiers within any network should be maintained. Once an individual reference code/number has been designated to a monitoring point/instrument, this identifier should not be reused in future for a different point/instrument (even if the initial one is removed).

5.4.3 Geographical location within the site

The geographical location and altitude of a monitoring point/instrument should either be recorded in the same manner as that recommended for sites (see **5.3.2**) or using its relative position to a specified site benchmark.

5.4.4 Operating period

The period of operation for any one monitoring point/instrument should be recorded, for example, the date of installation of a raingauge or removal of a stage measurement sensor.

5.4.5 Monitoring point/equipment details

Information relating to the measurement equipment used to record the data should be recorded. This should include, but not be limited to, the:

a) type of monitoring point/equipment;

b) details of manufacture/installation and setup; and

c) details of the resolution and uncertainty.

5.4.6 Operational details

Information relating to the operational practices used during the recording of the data should be recorded. This should include, but not be limited to:

a) the regularity of measurement; and

b) how readings are measured.

5.5 Observation metadata

5.5.1 General

In addition to metadata which relate to any one monitoring station, site or monitoring point/instrument within a site, information may be associated with a specific dataset or period of record within a dataset. Such metadata usually comprise two elements:

a) dataset metadata: information which relates to a whole dataset (see **5.5.2**); and

b) data flags/comments: information which relates to a specific measurement or period of measurement within the larger dataset (see **5.5.3**).

There should be a clear and unambiguous link from each observational metadata record/dataset and the monitoring point/site/station to which it relates.

5.5.2 Dataset metadata

Each monitoring station/site/point may have more than one time series of data associated with it (including both raw and derived data). There should be a clear and unambiguous link between each dataset and the monitoring point/instrument to which it relates, using the appropriate unique identifiers.

At each station/site/monitoring point, the primary time series for each variable at any one time should be clearly identified.

A catalogue of each dataset measured at a station should be included in the station metadata so there is a clear link between the station and any data associated with it.

Key information relating to each dataset should be recorded. This should include, but not be limited to, the:

a) unique series identifier;

b) variable being measured;

c) units;

d) interval;

e) period of record; and

f) method of derivation and details of the raw data series used (for derived series).

5.5.3 Data flags

Where there is a need to communicate information which relates to an individual measurement in time or a period of measurements within a dataset, this is best achieved through the use of data qualification flags.

One or more data flags should be recorded against each data point. They should be stored with the data to which they relate.

Data management systems should include clearly defined rules and scenarios outlining when each flag is to be applied.

If more than one flag is recorded against a single data point, a primary flag should be identified providing an overall indicator for the measurement. As a result, each data point should have a single, primary flag indicating one of the following:

U Unchecked: data have not yet been subject to quality control (either manual or automated) and an assessment of the most appropriate flag has not yet been made;

M Missing: value for the point in time is not available;

N Normal: measurement considered normal;

E Estimated: value is estimated (usually with additional secondary flags or metadata explaining how the estimate was made);

S Suspect: value is considered to be suspect (usually with additional secondary flags or metadata explaining why);

I Incorrect: value is considered incorrect (usually with additional secondary flags or metadata explaining why);

R Removed: value has been removed from the record (usually with additional secondary flags or metadata explaining why).

5.5.4 Data comments

In addition to data flags, additional metadata in the form of text comments might be required to provide information relating to one or more data points in a dataset, for example, a comment explaining that a record was affected by maintenance being carried out at a site between two dates. Such information should be linked to the dataset and time period to which it relates in an unambiguous fashion, and a summary of the most important points should be included in the site's summary description (see **5.2.5**).

Additionally, as recommended in **4.8.2**, where changes are made to datasets, details of these changes should be stored in metadata associated with the individual value/period altered. The metadata should include details of the change made (e.g. previous value/new version), who made the change and when, and why the change was made.

5.6 Hardcopy metadata

Key hardcopy metadata and other information should be digitized for inclusion in data management systems. Where this is not practicable, such information should be catalogued and stored in an information management system with an explicit link between each item and the monitoring site/data to which it relates.

6 Precipitation data

6.1 Raw data

6.1.1 Data to be recorded

With each measurement of precipitation, the following information, as a minimum, should be recorded in accordance with BS 7843 (all parts) and stored in its raw format.

a) Date and time of the end of the measurement period (in accordance with **4.5.1**).

b) Accumulation period covered by the measurement.

c) Measured precipitation amount in mm to the nearest 0.1 mm, including the water equivalent of new snow.

NOTE Some automatic raingauges have a resolution greater than 0.1 mm.

d) The value of the estimate of precipitation, if provided by the quality control process (see **6.5**).

6.1.2 Additional raw data for manual observations

The following supplementary information should be provided and archived electronically.

a) Occurrence of trace precipitation.

b) Water equivalent of new snow in millimetres (mm) to the nearest 0.1 mm.

c) If the observer indicates that the liquid collected is dew or melted frost, this should be noted along with any other comments on the nature of the precipitation.

d) Notes that might be helpful for the quality control of precipitation data (e.g. confirming that a thunderstorm caused an unusually large accumulation of rain).

Where data are provided from the local record:

1) that indicate missing observations;

2) for which the time of reading is not 09:00 UTC; or

3) that contain relevant weather remarks,

these should be retained within the data management system.

Where an observation or sequence of observations is missed, the data management system should have the capability of indicating that the next measurement is an accumulation.

For monthly accumulations, so that any necessary cross-month apportionments can be accounted for, the date and time of reading should be recorded if it was not 09:00 UTC on the first of the month.

6.1.3 Data storage

The data management system should be capable of accommodating:

a) time series of continuous or regular rainfall accumulations; and

b) sporadic or single-point-in-time rainfall observations, sometimes referred to as rainfall "event data".

6.2 Derived data

For daily measurements referring to the accumulated rainfall over the 24 hour period from 09:00 up to 09:00 UTC, each observation should be associated with the hydrological day (see **4.5.1**) and recorded against the start point of the accumulation (i.e. the calendar date of the preceding day).

The 24-hour total of sub-daily measurements from 09:00 to 09:00 UTC should be associated with the hydrological day (see **4.5.1**) and given the calendar date of the preceding day.

6.3 Resolution

The data management system should not degrade the measurement, and both daily and sub-daily rainfall should be archived to a resolution of at least 0.1 mm.

The data management system should be capable of storing readings below 0.05 mm as a trace (TR).

6.4 Data processing and formatting

6.4.1 General

The systems and processes should ensure timely receipt of rainfall data.

The data management system should be capable of identifying and retrieving missing data (for example, automatic observing systems should be designed to store data on site during periods of communication outage and to resend non-transmitted data when communication systems are restored).

Precipitation data, including precipitation metadata (see **6.6**), should be formatted for transfer between the observing site and the data management system. Processes for regular data exchange, ideally in real time, should be established.

6.4.2 Processing and formatting of manual observations

Observations from storage raingauges should be acquired for processing and storage as soon as possible after the measurement has been made and, for monthly paper records, as soon as is possible after the end of the month.

Manual observations should be transferred to electronic format by a process that minimizes the possibility of error. Where data on paper or rainfall postcards are converted to electronic format at a later date, dual keying, or other methods that ensure accurate transcription, should be used.

6.5 Quality control

6.5.1 General

Quality control should be conducted in accordance with **4.7.2** and, for precipitation data, the additional recommendations of **6.5.2** to **6.5.9**.

6.5.2 Principles of quality control

Quality control of precipitation data should be undertaken at various stages in the life cycle of a rainfall measurement. To minimize the time for which erroneous data might be held within a data storage system, checking procedures should begin at the earliest opportunity following their production.

The quality control process should identify erroneous measurements of precipitation while distinguishing them from extreme precipitation events and, where required, provide accurate estimates for missing or erroneous measurements which should be based on all relevant information available, including data from other sources (see **6.5.4.2**).

The quality control process should utilize all available relevant evidence, acknowledging the known variability of the weather and the likelihood of extreme precipitation events, and be applied consistently over time.

Errors in the quality control process should not be permitted to introduce biases to long-term climate averages, in particular to the frequency of extreme precipitation events.

Wherever possible, quality control checks should be applied consistently from raingauge to raingauge and over time. The methods of quality control and the dates of any changes to these methods should be recorded and maintained.

If new information comes to hand after routine quality control has been completed, procedures should allow for this information to be properly considered, and, if necessary, for repetition of some of the elements of quality control. Particular attention should be given to values which lie outside the range of values that have previously been experienced.

NOTE All quality control processes described in this and following subclauses are suitable for data contained in a national archive. They relate to measurements of precipitation comprising rainfall plus the water equivalent of new snow.

For most purposes, quality control should include a minimum of the primary and secondary levels given in **6.5.3.2** and **6.5.3.3**, respectively.

6.5.3 Levels of quality control

6.5.3.1 Level 1: No quality control

Where data have not been subject to quality control they should be flagged as unchecked (see **5.5.3**).

6.5.3.2 Level 2: Primary quality control

NOTE Sometimes referred to as "initial" or "real time" quality control.

Range and internal consistency checks should be performed in real time or near real time on sub-daily data being used operationally following the processes in **6.5.4.1**.

6.5.3.3 Level 3: Secondary quality control

NOTE Sometimes referred to as delayed mode quality control.

Using spatial analysis techniques, quality control of daily and monthly totals should be undertaken in delayed mode at a time determined by data availability and following the processes in **6.5.3.2**.

6.5.3.4 Level 4: Further quality control

NOTE Sometimes referred to as "post-delayed" mode quality control.

Further quality control checks should be implemented according to a regular schedule in order to identify problems such as drifting or step changes which would not be apparent from either initial or secondary quality control checks (see **6.5.4.3**).

6.5.4 Methods for quality control

6.5.4.1 Level 2: Primary quality control methods

Primary quality control checks should be applied to high time-resolution data at source and from within the destination data storage system in real or near real time. These checks should include the following:

a) a check for precipitation intensity that exceeds climatological extremes; and

NOTE 1 For a tipping bucket raingauge a minimum time between successive tips may be specified to identify cases where the bucket mechanism double bounces.

b) a check for isolated precipitation intensity peaks, e.g. very short period intense events that are not preceded or followed by precipitation of lower intensity.

NOTE 2 Such an event might indicate that the engineer has failed to isolate the system during testing.

Further checks should also be undertaken to identify blocked or faulty gauges by identifying spurious values.

Because of the large spatial and temporal variability of precipitation, quality control of short period precipitation using neighbouring measurements at distant locations is difficult to achieve successfully. Daily total values should therefore be derived from these high time-resolution data from automatic raingauges and presented for inclusion within the secondary quality control spatial analysis, for comparison with corresponding storage gauge data.

6.5.4.2 Level 3: Secondary quality control methods

Secondary quality control checks should be applied to daily and monthly totals from within the destination data storage system in delayed mode.

The principal method for the secondary quality control of precipitation data should employ spatial analysis techniques and be based on the comparison of measurements with those from neighbouring raingauges. The timing of this process should be determined by the optimum availability of as dense a network of comparison data as possible, and further defined by user requirements.

Statistical tests should be applied to assess the significance of the difference between a measured value and the value expected. Allowance should always be made for the large variability in daily precipitation accumulations that can occur between neighbouring locations on days dominated by deep convection. Allowance should also be made for instances where uncertainties in the estimates are large. For example, estimates are likely to be unreliable in mountainous areas where there are few neighbouring sites and local effects of the topography are significant.

The monthly totals from monthly gauges should be included for comparison with the monthly totals from all nearby raingauges.

NOTE 1 Precipitation measurement is highly dependent on the altitude of the raingauge site and orographic effects. For quality control purposes, site-to-site variability of precipitation may be reduced by normalizing with respect to the annual average rainfall at each site. Rainfall averages are normally calculated from 30 years' data, e.g. 1971-2000 (see BS 7843-4).

The information to be used to determine the uncertainty of a reading should include:

a) measurements from collocated raingauges at the same site; and

b) measurements from raingauges at neighbouring sites.

Additional factors to be taken into consideration include:

1) distance of each neighbouring raingauge from the site of interest;

2) distance of each neighbouring raingauge from other neighbours used in the estimate, favouring a spatially uniform distribution of neighbours;

3) altitude above sea level of each of the raingauges;

4) annual average rainfall at each of the comparison sites; and

5) uncertainty of each of the measurements used.

Further quality control evidence should be obtained from:

i) weather radar imagery;

ii) thunderstorm location systems;

iii) the synoptic weather characteristics during the measurement period;

iv) known characteristics of precipitation measurements at the site;

v) known characteristics of the raingauge, its exposure and the method of measurement;

vi) any additional information provided by the observer; and

vii) the history of the quality of measurements from the raingauge of interest.

Where a measurement from a raingauge spans two or more days, quality control methods should be applied using data from other sources to produce accurate reapportionments in order to reconcile any missing daily readings.

NOTE 2 Where a monthly gauge is not read at 0900 UTC on 1st of month, apportionment using data from near neighbours might be required for the intervening days.

6.5.4.3 Level 4: Further quality control checks

NOTE Quality control and monitoring checks are often applied to data on sub-daily, daily and monthly timescales to identify single instances of error. However, such checks are often unable to identify problems which on any one occasion are of a lesser magnitude, but which are highly significant when viewed over a longer period. For example, a measurement that is 25% too low due to underexposure of the raingauge by vegetation is unlikely to be flagged as erroneous on a single occasion, but the accumulated measurements from the raingauge over a period of a month or more might lie outside the range of values that might be reasonably expected.

Systematic errors in the measurements from a raingauge should be identified by the following methods.

a) Analysing departures from long-term averages using statistical techniques to test for significance.

b) Where a site has two or more collocated raingauges, analysing differences between the two sets of data over a period of time. Comparisons should be made with data from trials conducted under controlled conditions.

c) Analysing differences between raingauge measurements and radar estimates over a period of time.

d) Compiling summaries of the frequency of quality control flags raised for each raingauge. Those sites with a frequency significantly different from the average should be identified.

As soon as problems are identified appropriate corrective action should be initiated. The results of these checks should be recorded and maintained.

6.5.5 Late data

To capture any data which have been received late, a subsequent rerun or multiple reruns will be required. Where several quality control runs are performed, each subsequent run having a better coverage of data than its predecessor, previous flags should be replaced by flags from the latest run.

6.5.6 Quality control of frozen precipitation

Special care should be taken when dealing with snowfall measurements from tipping bucket raingauges. Manual measurements should be taken to ensure that snow is melted and recorded against the day it falls; automatic measurements might continue to show snow melt for some days afterwards, depending on the presence of raingauge heaters and the air temperature. Incorrect measurements are also produced by drifting snow accumulating in automatic raingauges after the actual time of snowfall. Manual quality control should be conducted in such cases, with values flagged as suspect, or, in the worst cases, whole periods being classed as "raingauge unserviceable" because of snowfall.

NOTE Other supporting measurements (for example, weather radar, temperature, snow depths, human observations, knowledge of the synoptic situation) can be very useful for quality control of these data.

6.5.7 Quality control flags

In accordance with **5.5.3**, quality flags should be applied to data that lie outside the range of values that might be reasonably expected. In doing so, the methods used should take account of the extent of atmospheric variability and give the benefit of any doubt to the measurement.

Where applicable, flags should be applied to both the sub-daily data and the associated daily value.

Other flags that indicate the possibility of errors with lower probability or confidence may also be applied to the data.

6.5.8 Quality control data comments

In accordance with **4.8.2** and **5.5.4**, reasons for making any changes to an observed value should be recorded. The reason for change, e.g. unindicated accumulation or tipping bucket raingauge blockage, should be recorded to form part of a history of quality of measurements from each raingauge. The quality control comments stored in the data management system should be used to assist with future quality control decisions by providing information on previous performance issues, some of which could be ongoing.

6.6 Precipitation specific metadata

6.6.1 General

In addition to the recommendations of Clause **5**, precipitation metadata records should be generated and managed in accordance with the following recommendations and those of **6.6.2** to **6.6.5**.

Precipitation metadata records should be kept permanently and organized in such a way that they can be linked to any measurement to which they relate.

Users of the data should be made aware of potentially poor quality data in snow events where a tipping bucket raingauge is not designed to operate in snow conditions.

NOTE Automatic weighing raingauges, however, make continuous measurements based on recording the weight of a vessel as rainfall accumulates. The weighing raingauge has the capability to record both liquid and water equivalent of solid precipitation.

6.6.2 Additional monitoring station metadata

Precipitation metadata relating to the station should include all elements given in **5.2**.

The geographical location of the station should be determined with respect to the primary raingauge.

The horizontal location should refer to the plot on which the measurement is taken and be accurate to within 10 m.

The elevation of the station is defined as the altitude above mean sea level of the ground on which the primary raingauge stands and should be accurate to within 1 m.

The station details should be applicable to all measuring devices at the location, which might include more than one raingauge and devices for measuring parameters other than precipitation. Where changes in the location of a raingauge occur at a station, details should be recorded by the raingauge deployment and communicated to users and stakeholders.

6.6.3 Additional monitoring site metadata

Precipitation metadata relating to the site should include all elements given in **5.3**.

6.6.4 Additional monitoring point metadata

In addition to the elements in **5.4**, the following metadata relating to each raingauge deployed at the site and its supporting infrastructure where applicable should be included:

a) raingauge type, i.e. manual or automatic;

b) raingauge model;

c) status of raingauge, i.e. primary or check;

d) geographical details (altitude of the base of the raingauge above mean sea level, height of rim above ground);

e) exposure assessment and deployment details (use of turf wall, pit or other sheltering device);

f) calibration details (method used, date of calibration and corrections if applicable);

g) maintenance, inspection and incident records (status, incident types and dates of visit);

h) observing and transmission system details (e.g. software systems that log and process data before transmission to users or archive and dates of change of system); and

i) observer details (if applicable, e.g. identification, training, dates).

6.6.5 Additional observation metadata

In addition to the elements in **5.5**, the following metadata relating to each raingauge dataset should be included:

a) precipitation measurement characteristics (e.g. 0.2 mm tip times, 24 h accumulations);

b) observing schedule (time the measurements are made, e.g. 09:00 daily readings, tip times, 1 min data);

c) from and to dates for the observing schedule;

d) reporting schedule (time data are sent from the site, e.g. monthly and real time);

e) from and to dates for the reporting schedule;

f) quality control methods; and

g) methods for correcting for systematic errors.

7 Water level data

7.1 Raw data

7.1.1 Data to be recorded

NOTE Raw data, in the context of water level monitoring, constitutes the original field measurements, their associated date/time-stamp and any associated metadata.

With each observation of water level, the following information, as a minimum, should be recorded in accordance with BS EN ISO 18365, and stored in its raw format:

a) date and time of the measurement (in accordance with **4.5.1**); and

b) measured water level in millimetres.

Every water level observation should be recorded with the associated date and time of the observation (see **4.5.1**). The level measured in accordance with BS EN ISO 18365 should be recorded as a time series variable applying to a period of time. In most cases, the time period should be from the start of recording to the present, but when the level of the zero (datum) does change (usually because the reference gauge is moved), the date/time of the change should be recorded.

7.1.2 Types of level data

Water levels may be recorded in lakes, reservoirs, rivers, canals, tidal waters, sewers, tanks, wells and boreholes. All measurements should be undertaken in accordance with BS EN ISO 18365. Water level data may be collected because level itself is the target information requirement, e.g. for groundwater or tides.

*NOTE 1 Where water levels are recorded relative to a local gauge datum, for example, as the height above a reservoir spillway, to subsequently calculate flows or volume, they are "stage data" and form the original time series from which a derived series of flow data is calculated, together with associated further derivations and statistics. The principles of **7.1.3** to **7.3.2** are applicable to stage data.*

*NOTE 2 Most aspects of data management relating to water levels are common to both surface and groundwater data; subclause **7.1.6** highlights points of particular relevance to groundwater levels collected in wells and boreholes, and to levels recorded in large reservoirs.*

7.1.3 Regular measurements and event data

Although a continuous record is often useful in validation, both on site and later during processing, regularly recorded intermittent data are sufficient for archiving and analysis. The choice of recording interval should be made with reference to the expected variability of the level with time.

NOTE Appropriate intervals are given in Table 1.

Table 1 **Recording intervals**

Locations	Interval
Storm sewer	1 min
Reservoir	1 h
River	15 min

Water level data are generally recorded at regular intervals using solid-state logging systems, although, historically, punched paper tape and magnetic tape have been used. Data may be regularly telemetered directly to computer. The observer's additional information supporting the raw data should be properly recorded and remain associated with the data file concerned. The data management system should include a system of including any additional notes (see **4.8**) and entering them onto computer when the data are processed.

When used for water level monitoring, event recorders record the time and water level only when the water level changes by more than a certain amount from its previous value, e.g. 1 mm. They perform the same function, in real time, as the chart digitizing operator; they capture only significant change points. As with chart processing, interpolation to regular intervals may be performed later.

7.1.4 Additional raw data for continuous measurements

NOTE An autographic chart is the only true recording device for continuous measurement as it is the only sure way of recording all high-frequency changes. Although charts have the advantage that important details can be recorded directly onto them by site staff, their usage has declined markedly in recent years due to the associated overheads of both storage and conversion to a digital format that can be used for the automated generation of derived data.

Where used, each autographic chart should include the following information:

a) the name of the site with sufficient detail to identify it uniquely (see **5.2.4**);

b) the type of sensor (for example, float, pressure-related) (see **5.4.5**);

c) the station identifying number (see **5.2.2**); and

d) the scales used for levels and time (see **5.4.6**),

and, for start and finish times:

1) the name of the observer;

2) the calendar date;

3) the time corresponding to the first (or last) appearance of the trace on the chart; and

4) the water level recorded by internal and/or external check gauges.

Continuous autographic data may be converted to digital form by digitizing the trace and storing in one of two ways:

a) as change points where the date, time and new level (whenever changed from the previous level by more than a specified value, usually 1 mm) are stored; or

b) by interpolation to regular intervals.

The digitizing process should reproduce the raw data (being the first digital "raw" record) and so follow the autographic trace precisely, though this might not always be necessary and might sometimes be undesirable, e.g. where it is considered more useful to capture the underlying "signal" than the confusing "noise" (oscillations). The operator should include explanatory comments to fully describe the procedure followed. Any adjustment to the data to reflect real time or to match measured levels from a reference gauge should be conducted and flagged accordingly in a parallel time series. The digitizing process should also allow for all observer comments (site metadata) to be entered and locked in with the data.

7.1.5 Resolution and uncertainty

Water levels should be recorded to a resolution of 1 mm.

7.1.6 Groundwater levels and water levels in reservoirs

Groundwater levels are commonly measured as the distance *below* a datum point to the water level. Whilst this could have local significance, these data should be converted to a level *above* the national datum, either as a function within the field data logger or within the data management system.

NOTE Thus, as the water level rises, so too does the recorded value (which would otherwise increase with falling level due to the increased measured distance from the reference point), and levels at a point can be used to generate regional information.

The same principle should be applied, where necessary, to the levels recorded at large reservoirs, where the reference point for measurement is frequently the level of the reservoir spillway.

The data management system should be capable of indicating whether measurements are recorded above or below the station datum, including any such references in the station metadata.

7.2 Quality control

7.2.1 General

Quality control should be conducted in accordance with **4.7.2** and, for water level data, the additional recommendations of **7.2.2**.

As indicated in **4.8.2**, quality control should be conducted such that any changes required to the data are made not to the raw data but within a parallel data series.

7.2.2 Levels of quality control

Where data have not been subject to quality control (Level 1) they should be flagged as unchecked (see **5.5.3**). Each of the methods in **7.2.3** may be conducted to increasing degrees for each of the quality control levels 2 to 4.

7.2.3 Methods for quality control

7.2.3.1 Use of reference gauges for revision of level data

Manually-recorded readings from reference gauges should be compared with automatically recorded readings made at the same time. Automated readings should be corrected to agree with manual observations. The data management system should record whether such corrections have been made, the time period to which any changes apply, and the means by which the changes were made.

As a result of the validation process, some constant addition or scaling may be applied over a known period of time. Such changes should be recorded and therefore made reversible. Other changes (for example, infilling, removal of spurious spikes, estimations during periods of blockage or freezing) may not be considered reversible.

All revisions should be stored in the database with comments intimately attached.

7.2.3.2 Corrections for timing errors

Date and time information supplied by an external control (a telemetry base station, the observer or data processing staff) should be used to correct the times given by the internal clock of the recording instrument. Unless there is evidence to the contrary, it should be assumed that any error in instrumentation clock speed is constant over the period of observation.

The date and time of each data transfer from field instrumentation to data management system should be logged to permit monitoring of routine data transfers and ensure continuity between successive batches of data.

7.2.3.3 Visual inspection of record

Visual checking of a graphical record of water level (hydrograph) should be conducted by experienced hydrometric staff as part of data validation. Hydrographs should be inspected on original charts or after generation from the data management system, which should be capable of plotting either as hardcopy or on screen.

Hydrograph patterns are useful in the diagnosis of a variety of instrument malfunctions and site problems. They also provide graphic illustration of the effects of artificial disturbances generated upstream. Characteristic patterns should be recognizable at each site, and anything unusual should be investigated, and, if necessary, comments attached to the data.

An experienced operative should scan the data on screen whilst at the same time checking any chart record, together with any notes and comments brought back from the field. These checks should be completed before any correction of levels or times (see **7.2.3.1** and **7.2.3.2**), and should incorporate options for making corrections. If the only records are in the form of charts, these validation procedures should be included in the digitizing process.

In some cases, such as when two level records are associated with a site (for example, upstream and downstream of a control structure or measurement reach affected by variable backwater), the hydrographs should be compared on screen. The data management system should allow this.

7.2.3.4 Arithmetic checks

Preliminary validation should include an automatic check against preset values describing the maximum and minimum expected levels and the maximum and minimum expected rates of change. The limits should be set initially at 90% of the maximum or minimum observed values from the preceding period of record, and should be revised according to experience. When a limit is passed, the data processing operator should be required to confirm that the particular sequence has been examined, that the record is valid, and that the validation limits have been adjusted if necessary.

There should be automatic checks of continuity in level between contiguous batches of water level data.

7.2.4 Data estimation and infilling

Although every effort should be made to produce a sensibly continuous record of level, infilling should be conducted only when there is good evidence to support it. Gaps in water level data should be infilled in accordance with **4.7.3** and the following recommendations.

Irregular time series, by definition, contain gaps, but once a regular time series of data has been commenced, this should be maintained until a decision is made to close the record. The method used to infill any gap in records should be selected according to the size of the gap, the circumstances that caused it, and the availability of alternative data, but may include (in order of preference):

a) use of data from a parallel backup instrument at the same site;

b) use of data from an external source, whether field observations or modelled from another monitoring site in the same, or nearby, catchment with similar characteristics;

c) automated interpolation between the known fixed points at the start and end of the gap, which may be straight line or spline-fit, according to the shape of the known portions of the hydrograph; and

d) manual insertion, which should only be undertaken by a trained operator and, wherever possible, be based upon external data, e.g. insertion of a missing peak based on surveyed trash marks following a flood.

As recommended in **5.5.3**, the data management system should allow flags to be allocated against any such data insertions, whether individual data points or multiples thereof. All such data should be flagged as "estimated" as a minimum and preferably with a percentage confidence that the estimation is within 5% of the true value that would otherwise have been recorded.

Details of the infilling techniques used should be recorded in the metadata. Where possible this should include an indication of the methods' uncertainties and details of any external sources of data used.

7.3 Water level specific metadata

*NOTE For water levels there are generally no monitoring station or observation metadata to be recorded in addition to those recommended in **5.2** and **5.5**, respectively.*

7.3.1 Additional monitoring site metadata

In addition to the elements in **5.3**, the level of reference gauge zero and other significant levels in the instrument house, and a clear indication of the datum used, should be recorded as recommended in **5.3.2**. The monitoring station details should be held as part of the data management system, rather than held in hard copy files.

*NOTE Further details are needed for a regular (flow) gauging station (see **8.7**).*

7.3.2 Additional monitoring point metadata

Water level metadata relating to the station should include all elements given in **5.4**. Particular attention should be given to ensuring that the vertical level of gauge zero for all water level monitoring points is recorded in accordance with **5.4.3**.

8 Velocity and discharge data

8.1 Raw data

8.1.1 Data to be recorded

With each measurement of discharge, the following information, as a minimum, should be recorded in accordance with BS EN ISO 18365. All parameters should be stored as raw or original, i.e. as collected in the field, including:

a) date and time of the measurement (in accordance with **4.5.1**);

b) stage (see Note 1 to **7.1.2**);

c) flow (if derived on site);

d) velocity (if measured on site); and

e) cross-sectional area (if derived on site).

8.1.2 Additional raw data for irregular/spot measurements

Irregular and "on/off" discharge measurements may be estimated at a wide range of sites and are usually referred to under the general heading of "spot gaugings".

The original gauging records should be retained.

Whatever technique is used to measure or estimate the discharge, in addition to the information recommended in **8.1.1**, the following details supporting the discharge values should be stored:

a) location, including grid reference;

b) the start point (the river bank);

c) the method (e.g. cableway, bridge, wading, boat);

d) the depth sampling arrangement and number of sections;

e) the instrument type and reference number, and the date of the most recent calibration [acoustic Doppler velocity profilers (ADVP), current meter];

f) textual description of the conditions; and

g) name of observer.

8.1.3 Additional raw data for regular or fixed site measurements

Direct measurement of discharge at sites, and the regular estimation of discharge at fixed sites, can be made in a number of ways:

a) derived from water level – stage/discharge method;

b) ultrasonic time of flight systems;

c) velocity index method, which incorporates various measurement systems such as:

   1) side looking or bed-mounted ADVPs;

   2) range-gated devices; and

   3) radar velocity systems; and

d) electromagnetic systems.

The parameters given in Table 2 should be recorded when calculating discharge by the various methods.

NOTE Recording the parameters in Table 2 ensures that the flow can be recalculated if a problem with the flow is subsequently discovered.

Table 2 **Parameters to be archived when calculating discharge, as appropriate**

Stage discharge method	Ultrasonic time of flight systems	Velocity index method	Electromagnetic systems
Date	Date	Date	Date
Time	Time	Time	Time
Stage	Stage	Stage	Stage
Weir dimensions and constants	Discharge	Discharge: theoretical and/or indexed	Discharge
Drowned flow correction; crest tappings	Velocities – mean and individual paths	Measured velocity	Mean velocity
Rating equation	–	Index rating	–
Gate positions when used in flow calculations	Temperature	Temperature	Voltage
–	–	–	Coil current

8.2 Data processing

NOTE Once raw data have been collected, they have to be processed and archived. Depending on the method of flow calculation this will involve further calculation and processing which requires the storage of additional parameters. This subclause sets out the specific requirements for each method.

8.2.1 Discharge data derived from water levels

8.2.1.1 General

The stage-discharge relationship is the relationship at a gauging station between stage and discharge, and is sometimes referred to as a rating curve or rating. The methods of determining a stage discharge relationship should be as described in BS ISO 1100-2. [1)]

8.2.1.2 Rating curve

All computed discharges should be unambiguously associated with a specific rating curve.

Full details of the rating curve should be archived including regression coefficient or standard error.

All gaugings used to create a rating curve should be archived, in accordance with **8.1.2**.

Theoretical and fitted relationships should be entered in the database, e.g. weir formulae.

All relationships should be given a reference number and a date range.

A clear history and record of all ratings applied should match the full period of record of the parameter and/or site.

All working iterations should be deleted once a rating has been activated.

All historic and superseded rating equations should be archived, along with the flow measurements from which they were derived and the period of time for which they were applicable.

[1)] A revision to BS ISO 1100-2 is in preparation, which will eventually be published as BS ISO 18320.

8.2.1.3 Data management problems specific to the stage discharge method

As the flow is derived using the basic formula of Velocity × Area any factors impacting on either of these will result in a variation of the calculated flow. When managing discharge data, the various factors that can affect the calculation of flow when using the stage discharge method should be considered, including:

a) hysteresis, where flow for a given stage varies between the rising and falling hydrograph;

b) errors in stage reading [stilling well lag; stilling well versus in-river differences (velocity drawdown)];

c) sensitivity of station;

d) drowning of hydraulic structure;

e) flow control impacted by a change in downstream conditions; and

f) fouling, obstructions, sediment and weed growth in gauging section.

The data should be corrected to allow for any of these factors or the factors should be recorded in the metadata.

8.2.1.4 Use of multiple observations of stage

Where multiple stages are used to calculate the flow, i.e. crest or tail water used for drowned flow correction, then this should be clearly recorded in the metadata.

The method used to calculate the flow if it is being modified by a crest or tail water correction should be recorded. The amount of correction or reduction factor applied should also be recorded, and any flow data that have been modified should be flagged as such.

8.2.1.5 Reprocessing of stage discharge data

Reprocessing of historic data can be required when new additional data have been gathered that allow a better stage discharge to be calculated, e.g. additional high flow gaugings. Any reprocessing of data should be recorded in the metadata, indicating exactly what was done.

Reprocessing of flows over the period of record following change in the rating curve may be limited to the period of the rating change or may apply over historic data. If this is the case, the rationale for the change should be recorded in the metadata.

*NOTE Recalculation of the index rating might be necessary for similar reasons as stage discharge relationships. If the index is recalibrated this might require the flow discharge also to be recalculated from the period of the changed index curve. If this is the case then the rationale for the change needs to be recorded (see Clause **5**).*

8.2.2 Discharge data from time of flight ultrasonic systems

For data management specific to time of flight ultrasonics, the following should be recorded:

a) the number of acoustic paths deployed;

b) the number of paths operating at any given time;

c) any change in the cross-sectional area; and

d) any obstruction or debris blocking the ultrasonics paths.

8.2.3 Discharge data from velocity index method

For data specific to index systems, the following should be recorded:

a) change in stage/area; and

b) position of the indexed instrument within the water column.

Reprocessing of data calculated using the velocity index method should be consistent with that for stage discharge in **8.2.1.5**.

8.2.4 Discharge data from electromagnetic systems

If water levels exceed the level of the bed liner, which is likely to cause a negative result on the flow which will be incorrect, this should be recorded.

Any electromagnetic interference detected should be recorded as this will have a detrimental impact on the flows recorded.

8.2.5 Combining discharge data from multiple methods

When combining the methods in **8.2.2** to **8.2.4**, to calculate the total flow at a monitoring site/station the flow for each monitoring point should be calculated, complete with any edits, corrections and flags. These data should then be combined to calculate a complete flow measurement for the location as a whole.

A similar method should also be applied when calculating total flow at a single location where multiple methods are used to measure the flow at the monitoring point. This becomes a hybrid flow site. In this case the flow from each method should be calculated, edited flagged as previously described and then combined with the other method, e.g. a low flow is calculated using an established stage discharge rating at higher levels the flow is calculated using a side-looking Doppler system which has been indexed.

8.3 Resolution and uncertainty

Discharge data should be stored to four significant figures or three decimal places, whichever represents the larger flow. Data for particularly small measurements may be stored to a higher resolution if required and if the method of measurement warrants confidence in the increased resolution.

8.4 Derived data

8.4.1 Daily mean flow

The daily mean flow should be calculated in accordance with **8.2**, from the sequence of samples that have been recorded over the day.

NOTE It is common practice to do this, although there are many catchments where the daily value is too coarse for most studies, and others for which the practice is unnecessarily detailed. The advantage of computing the daily mean discharge is that it provides for comparison between catchments and also relates usefully with daily rainfall data.

Daily mean flows should be expressed either in cubic metres per second or in megalitres per day. The units should be stated in the dataset metadata.

Daily mean discharge should be computed by integrating the volume below the continuous hydrograph over 24 h. When calculated from regular interval discharge data recorded every 15 minutes (as is common place in the UK) a straight averaging of the ninety-six consecutive values within the 24 hour period should be used.

The boundary between one day and the next should be clearly stated in the dataset metadata. This should normally be the hydrological data starting at 09:00 UTC (see **4.5.1**).

8.4.2 Peak discharges

Annual maxima and peak-over-threshold (POT) time series should have independence criteria to ensure that valid peaks are independent between each other and water years.

8.5 Naturalization of discharge data

Where necessary for hydrological analysis, artificial influences on the gauged flow record should be removed. Examples include assessment of natural catchment yield, comparisons of hydrological characteristics between catchments, and investigation of natural long-term trends.

To permit flow naturalization, data relating to significant point discharges to and abstractions from the river upstream of the gauging station or reservoir storage should be collected. These data should be separately processed and validated as necessary. If the data have to be obtained from other organizations, then care should be taken to ensure that the data quality is fully understood.

Naturalization should be based on daily mean flows wherever possible, but monthly calculations might be the only feasible option in some cases, e.g. when correcting for the effects of an upstream reservoir.

Selection of gauged discharge records for routine naturalization should be determined by the importance of the station to long-term monitoring in the region and by the cost of acquiring and validating all the necessary data. A hydrometric agency should ensure that the data gathered from a significant subset of its gauging stations are either natural or naturalized. Data can be regarded as natural if the maximum disturbance to natural flow is less than 10% at or below Q_{95}.

All data should carry explicit flags or comments to define either the extent of upstream influences affecting the gauged record or details of the naturalizing process. This should be done even when the gauged record is acceptably natural.

8.6 Quality control

8.6.1 Levels of quality control

The procedures for the quality control of hydrometric data should be structured around the four levels given in **4.7.2**. Recommendations for what procedures should be included in each quality control level for discharge data are given in **8.6.2**.

8.6.2 Methods for quality control

8.6.2.1 Level 1: No quality control

Data should be stored, but marked unchecked (see **5.5.3**).

8.6.2.2 Level 2: Primary quality control

The following activities should be undertaken:

a) conduct a graphical and numerical check of the data to look for data anomalies;

b) investigate and resolve any issues with missing or spurious data and a reactive field visit if necessary; and

c) mark the data edited if any changes have been made.

8.6.2.3 Level 3: Secondary quality control

The following activities should be undertaken:

a) confirm Level 2 actions are complete;

b) check the sub-daily and derived daily mean flow records for the batch of data being validated with the previous 6 months;

c) plot sub-daily 15-minute and derived daily mean flow data ensuring all new data and at least the previous 6 months of data are displayed;

d) plot spot flow gaugings for the site onto flow hydrographs to assess the reliability of the time series data;

e) where records from backup instruments are available, check for differences between the time series plots, discontinuities and behaviour that is uncharacteristic for the site;

f) if the data are clearly erroneous, remove them and treat the resulting gap as missing data - be cautious before removing peak flows;

g) where no backup data are available or it is not practical to infill the sub-daily time series, it might be more appropriate to infill the daily mean; and

h) add comments as necessary, incorporating information from field logs if they are not already recorded with the data.

8.6.2.4 Level 4: Extended quality control

The following activities should be undertaken:

a) confirm Level 3 actions are complete;

b) compare flows for the last water year with the entire period of record for the site, examine new extremes and look for trends;

c) plot the entire period of record at a daily and sub-daily time step, where appropriate, during the first extended check only, thereafter limiting regular inspection to the last 15 to 20 years;

d) give special consideration to extremes that are at or beyond the limits of the structure or measuring instruments or, in the case of derived flows, beyond the confirmed range of the rating need special consideration, adding comments to as appropriate;

e) compare the record of flows with analogue sites, looking for unexplained discrepancies, e.g. where there is greater flow at the upstream site than at the downstream site; and

f) where precipitation data are available, check the catchment water balance for the last water year, comparing it with the entire period of record to identify changes.

8.6.3 Review of extreme flows

Extreme flow analysis should be undertaken using a range of statistical and modelling methods to ensure a reality check for extreme events.

Any verification by site visit post-event should be archived along with photographs and/or recordings of wrack marks.

Adjacent catchment comparisons should be undertaken.

8.6.4 Data estimation and infilling

Gaps in discharge data should be infilled in accordance with **4.7.3** and the following recommendation. Infilling of discharge data should be completed using one or more methods dependant on the characteristics.

Where discharges are calculated from stage records the recommendations in **7.2.4** should be applied. Under such circumstances preference should be given to infilling the source stage record. However, where this is not possible, the direct infilling of derived discharge data should be completed where sufficient supporting information is available.

The following procedures (in order of preference) should be followed for where gaps in discharge time series are identified.

a) Use of discharge data from a backup system/instrument(s) at the site to infill gaps in the main station record when the primary data source fails.

b) When no suitable backup data are available an appraisal should be made of what data it is appropriate to infill. Consideration should be given to the:

   1) likely range of discharge values to be infilled;

   2) size of the gap; and

   3) supporting evidence available.

   Particular attention should be given to appraising the time step at which it is appropriate to infill. For example, the information available might not be sufficient to infill a 15-minute resolution discharge time series to an appropriate level of uncertainty, but an estimate of the daily mean flow might be possible.

   In most circumstances, where sufficient supporting information and/or data are available to give confidence in the resulting estimates of discharge, the time series should be infilled.

   Preference should be given to infilling time series gaps completely. Where this is not possible and gaps in time series are likely to compromise the calculation of peak discharge, infilling of those peaks alone can be undertaken if there is sufficient supporting information (for example, wrack marks or gaugings).

c) Infilling of discharge data should be completed using one or more method dependent on the characteristics of the water body and the particular nature of the data gap. A wide range of methods exist, including those outlined in Table 3, which provides guidance on the selection of the most suitable technique.

   In all cases, the choice of infilling technique and its subsequent application should be made with help from staff who are familiar with the water body.

As recommended in **5.5.3**, the data management system should allow flags to be allocated against any such data insertions, whether individual data points or multiples thereof. All such data should be flagged as "estimated" and the justification for each occurrence recorded. Details of the infilling techniques used should be recorded in the metadata. Where possible this should include an indication of the methods' uncertainties and details of any external sources of data used.

When infilling gaps in discharge records, use should made of supporting information and/or data such as flows from a nearby (or analogous) water body which is hydrologically-similar and/or rainfall data which provides an indication of catchment inputs. Some of the techniques in Table 3 rely on the transfer of discharge information from other analogue stations and these should normally be up or downstream of the station concerned or be in a nearby catchment. The choice of analogue station should take account of geographical proximity, similarity of runoff generation mechanisms and catchment characteristics, or impacts of artificial influences which significantly affect the natural flow regime.

8.7 Discharge specific metadata

*NOTE For volume there are generally no station or site metadata to be recorded in addition to those recommended in **5.2** and **5.3**, respectively.*

8.7.1 Additional monitoring point metadata

In addition to the elements in **5.4**, the following should be recorded, where applicable:

a) maximum measurable flow of the structure/rating or system;

b) maximum flow that the measured channel can contain (bank full flow); and

c) level at which flow bypassing occurs.

Any estimates of flow bypassing should be recorded and, if possible, added into the complete station flow record. If it is not possible to correct the flow time series to accommodate for bypassing flow then suitable comments and flags should be added to the record.

8.7.2 Additional observation metadata

Secondary flags should, where appropriate, be used to provide additional information for:

a) instrument malfunction;

b) instrumentation drift;

c) debris, temporary obstructions, bed changes, siltation;

d) peak truncation;

e) stage-discharge rating and/or weir performance;

f) data that are the result of drowned flow corrections; and

g) flows that exceed the rating limit.

During computation of daily mean values from the sequence of discharge data, the processing program might encounter flags and/or comments relating to one or more values in a given day. This information should be carried over. Daily mean values should be flagged with completeness of the computation, i.e. complete/incomplete, and preferably a measure of that completeness, e.g. 80%. Relevant data comments should be available to the archived daily mean values.

Table 3 – Techniques for infilling discharge data

Method	Comments	Suitability
Manual inference	Estimates are derived manually based on available supporting information (for example, visual comparison with records from other locations). This is an intuitive estimate.	Accuracy can normally be assured for short gaps with no rainfall or longer gaps during stable recessions, but other conditions might lead to increased difficulty and subjectivity in determining estimates. This method can be used to correct displacements caused by work at a site, or instrument problems such as blockages, weed growth clearance, obstructions or lighting strikes.
Serial interpolation	Estimates are derived using linear, polynomial or spline interpolation. Such techniques close a gap in a time series by: • extending a trend between the recorded data points either side of the gap (for example, an exponential decay during low flows); • simple bridging using a straight line; • inserting a non-linear/curved line that can be used for inserting peaks or troughs, especially where additional information is available.	These techniques are likely to only be successful throughout stable periods and should be used to fill short gaps where estimates can be confirmed by supporting information. The discharge series should be expected to behave in a steady way over the period of the data gap and, therefore, the techniques should not to be used for highly variant time series. Interpolation should not be used if it causes a sudden step in the data that is untypical for the site.
Scaling factors	Estimates are derived by multiplying discharge data from other hydrologically-similar water bodies by a scaling factor (such as the ratio of the catchment area).	These techniques are largely unsuitable for infilling gaps in time series data, but might perform effectively in estimating summary hydrological characteristics (e.g. annual mean flows) in regions with largely homogeneous hydrological conditions.
Equipercentile techniques	The discharge percentiles at the station where flows are being estimated and those at one or more hydrologically-similar water bodies are assumed equal for any given point in time. Gaps in the target time series are infilled by calculating the flow percentile values for the similar water body and then using the existing target time series to derive the flows equivalent to these percentile values.	These techniques have broad applicability where records from hydrologically-similar water bodies exist.
Linear regression	A regression equation between time series being infilled and flows recorded at at least one other hydrologically-similar water body is derived, commonly via the least squared method, and used to estimate flows during the gap.	Improved estimates of discharge can be derived by using data from multiple other hydrologically-similar water bodies.
Hydrological modelling	Methods for deriving estimates vary from black-box modelling, whereby the model inputs are related to the outputs without considering the processes involved, to more complex process-based models and use of artificial neural networks.	These methods offer the potential to generate highly accurate estimates. However, such techniques are often too resource-intensive for rapid, widespread, application.

9 Volume data

9.1 General

Storage volumes in lakes and reservoirs should be computed from recorded water level data using theoretical or empirical relationships. Level data should be treated in accordance with Clause **7** and subclauses **9.2** to **9.4**.

9.2 Level/storage relationships

The relationship between measured level and volume in a surface storage should be obtained by hydrographic survey, by aerial survey at various stages of drawdown, or by other means as appropriate (see PD ISO/TR 11330). The relationship calculated between level and volume should be described and dated in the same way as a stage-discharge relationship for a gauging station. Such relationships should extend above the level of the natural control, or reservoir spillway, to at least the maximum level anticipated.

9.3 Frequency of sampling and processing

Levels may be sampled and processed to storage, and daily means computed in the same way as with discharge. However, a daily mean storage is of limited interest, so a record of storage should be generated at a particular time each day. This may be performed either by making manual observations at that time or by extracting a single value from a logged record (or interpolating from a digitized record).

9.4 Resolution and uncertainty

The resolution and uncertainty of volume data should be suitable for the total volume of storage.

9.5 Volume data specific metadata

NOTE For volume there are generally no station or observation metadata to be recorded in addition to those recommended in ***5.3*** *and* ***5.5****, respectively.*

9.5.1 Additional monitoring site metadata

In addition to the elements in **5.2**, the following volume data metadata relating to the site should be recorded:

a) the source of the level-storage relationships and their dates of derivation and applicability;

b) a description of outflow arrangements and capacities; and

c) the intended use of the storage (different uses generate different change patterns in storage, and in the flows of downstream watercourses, which can be used in data quality control).

9.5.2 Additional monitoring point metadata

In addition to the elements in **5.4**, the position of the level recorder or staff gauge relative to the storage body as a whole should be recorded.

Bibliography

Standards publications

For dated references, only the edition cited applies. For undated references, the latest edition of the referenced document (including any amendments) applies.

BS 7843-4, *Acquisition and management of meteorological precipitation data from a raingauge network – Part 4: Guide for the estimation of areal rainfall*

DD CEN ISO/TS 25377, *Hydrometric uncertainty guide*

ISO/IEC Guide 98-3, *Uncertainty of measurement – Part 3: Guide to the expression of uncertainty in measurement*

PD ISO/TR 11330, *Determination of volume of water and water level in lakes and reservoirs*

Other publications

[1] GREAT BRITAIN. The Public Records Act 1958. London: The Stationery Office.

[2] OPEN GEOSPATIAL CONSORTIUM (OGC). *WaterML2.0 Part 1: Time Series Encoding Standard*. OGC. 2014. (http://www.opengeospatial.org/standards/waterml) [2)]

[3] WORLD METEOROLOGICAL ORGANIZATION: *Manual on Codes – International Codes Volume I.1, Part A – Alphanumeric Codes*: 2010 WMO-No 306 Vol I.1.

[4] WORLD METEOROLOGICAL ORGANIZATION: *Manual on codes – International Codes Volume I.2, Part B – Binary Codes*: 2010 WMO-No 306 Vol I.2.

[5] GREAT BRITAIN. The Data Protection Act 1998. London: The Stationery Office.

2) Last accessed 16 December 2014.

This page deliberately left blank

British Standards Institution (BSI)

BSI is the national body responsible for preparing British Standards and other standards-related publications, information and services.

BSI is incorporated by Royal Charter. British Standards and other standardization products are published by BSI Standards Limited.

About us

We bring together business, industry, government, consumers, innovators and others to shape their combined experience and expertise into standards -based solutions.

The knowledge embodied in our standards has been carefully assembled in a dependable format and refined through our open consultation process. Organizations of all sizes and across all sectors choose standards to help them achieve their goals.

Information on standards

We can provide you with the knowledge that your organization needs to succeed. Find out more about British Standards by visiting our website at bsigroup.com/standards or contacting our Customer Services team or Knowledge Centre.

Buying standards

You can buy and download PDF versions of BSI publications, including British and adopted European and international standards, through our website at bsigroup.com/shop, where hard copies can also be purchased.

If you need international and foreign standards from other Standards Development Organizations, hard copies can be ordered from our Customer Services team.

Subscriptions

Our range of subscription services are designed to make using standards easier for you. For further information on our subscription products go to bsigroup.com/subscriptions.

With **British Standards Online (BSOL)** you'll have instant access to over 55,000 British and adopted European and international standards from your desktop. It's available 24/7 and is refreshed daily so you'll always be up to date.

You can keep in touch with standards developments and receive substantial discounts on the purchase price of standards, both in single copy and subscription format, by becoming a **BSI Subscribing Member**.

PLUS is an updating service exclusive to BSI Subscribing Members. You will automatically receive the latest hard copy of your standards when they're revised or replaced.

To find out more about becoming a BSI Subscribing Member and the benefits of membership, please visit bsigroup.com/shop.

With a **Multi-User Network Licence (MUNL)** you are able to host standards publications on your intranet. Licences can cover as few or as many users as you wish. With updates supplied as soon as they're available, you can be sure your documentation is current. For further information, email bsmusales@bsigroup.com.

Revisions

Our British Standards and other publications are updated by amendment or revision.

We continually improve the quality of our products and services to benefit your business. If you find an inaccuracy or ambiguity within a British Standard or other BSI publication please inform the Knowledge Centre.

Copyright

Useful Contacts:

Customer Services
Tel: +44 845 086 9001
Email (orders): orders@bsigroup.com
Email (enquiries): cservices@bsigroup.com

Subscriptions
Tel: +44 845 086 9001
Email: subscriptions@bsigroup.com

Knowledge Centre
Tel: +44 20 8996 7004
Email: knowledgecentre@bsigroup.com

Copyright & Licensing
Tel: +44 20 8996 7070
Email: copyright@bsigroup.com

BSI Group Headquarters

389 Chiswick High Road London W4 4AL UK

bsi.

...making excellence a habit.™