Hydrological data management systems within a national river flow archive

M.J. Fry and O. Swain

Centre for Ecology & Hydrology, Wallingford, UK
Email: mfry@ceh.ac.uk

Abstract

The UK National River Flow Archive (NRFA) is the UK's focal point for hydrometric data. It provides data and expertise on river flow conditions to the public, to government, and to research and commercial organisations. Quality control and improvement of data, through validation and resolution of monitoring and data issues, is an important focus of its work, with the aim of maximising the utility of the national archive for a wide range of end-users. Recent developments in the NRFA have led to innovative systems for data management, analysis and visualisation. These systems increase the capacity to improve data by integrating river flows with rainfall and other time series data, spatial information such as catchment land use, elevation and geology, and metadata providing information on gauging station quality. Flow data are disseminated alongside detailed metadata ensuring that the end-user can exploit the data effectively, and is aware of any limitations. This model for a hydrological data management system could potentially be applied in other hydrological networks and data archives, both in the UK and internationally.

Introduction

Hydrological data underpin the management of the water environment. Data are required for water resources assessment, flood risk management and for a range of regulatory and research activities. Statistics from river flow data underpin water policy initiatives and the drafting of protocols and legislation — for both national and international application. Hydrological archives are essential for the provision of data for these purposes. This paper describes the activities undertaken by the UK National River Flow Archive (NRFA) to ensure that long term river flow records are available to suit these needs, and the data management systems that allow this to be achieved.

The UK National River Flow Archive

Flow measurement in the UK is primarily undertaken by the Environment Agency for England and Wales (EA), the Scottish Environment Protection Agency (SEPA) and the Rivers Agency for Northern Ireland, here jointly referred to as the Measuring Authorities (MAs). River levels are monitored, generally at 15 minute intervals, and converted to flows within local or regional MA offices. Daily mean flows for some 1400 gauging stations are provided at regular intervals (monthly or annually) to the National River Flow Archive (NRFA) at the Centre for Ecology and Hydrology. The NRFA provides a secondary data validation and archive service, and forms the focal point for access to hydrometric data for the UK as a whole. The archive serves a wide user community, incorporating water management professionals, scientific researchers, educational users, national government bodies and international organisations. However, its remit goes beyond the validation, storage and provision of hydrological data. Marsh (2002) outlines the additional national capability required to be provided by the NRFA to meet UK strategic needs, including establishing benchmark hydrological conditions, and assessing and monitoring national water

resources, and changes in river flows. It also provides a national perspective on the utility of the UK gauging station network and the utility of archived UK river flow data. Whilst these requirements are specific to meet UK national capability, equivalent and comparable requirements are met by National Hydrological Services (NHS) around the world.

The NRFA meets these requirements through the following specific activities:

- Regular acquisition of daily river flow data from MAs;
- Annual validation of river flow time series, including quality control and the resolution of data quality issues;
- Maintenance of metadata providing quality assurance, describing a range of factors affecting the flow records from measurement techniques to changes within the catchment over time;
- Production of key catchment statistics and indicators of catchment utility and representativeness;
- Monthly and annual reporting on UK hydrological conditions, under the aegis of the National Hydrological Monitoring Programme;
- Occasional reporting on important hydrological events such as major floods and droughts;
- Production and publication of key hydrological statistics;
- Identification of a network of natural catchments and the assessment of trends in river flows in these catchments;
- Maintenance of a website and manual data retrieval service.

These activities require a detailed understanding of contemporary hydrological conditions, as well as a longer term perspective on historical river flows and a national-level overview of the ability of the UK gauging station network to capture hydrological conditions. This work is data intensive, and is dependent upon the integration of a wide range of hydrological information, including time series, such as rainfall and groundwater data, spatial information describing factors affecting hydrological processes, and gauging station information on the quality of measured flows. The ability

to extract relevant information from the large number of complex datasets available is key to achieving the core aims of the NRFA. IT systems are required for the storage and management of data, the integration and visualisation of data, and the production of a range of detailed statistics.

NRFA data management systems

The range of hydrological regimes across the world, and the different level of data availability between countries, means that hydrological systems requirements of NHSs vary from country to country. However, there are some core requirements, namely for the handling and processing of time series of river flows, that are common to all hydrological archives. Figure 1 shows a schematic representation of NRFA data processes from acquisition to dissemination. Numerous commercial packages exist to provide standard functionality for time series data management, including visualisation, conversion between intervals and the production of statistics. Several packages are specifically tailored for hydrological data, providing additional functionality, e.g. for rating development and management and flow duration calculations: WISKI (Kisters, 2010a), TIDEDA (NIWA, 2010), HydroLog (Isodaq, 2010), HYDATA (CEH, 2010), Hydstra (Kisters, 2010b), etc. These packages meet the principal requirements for hydrological data management systems, often including the retrieval of data from telemetry systems for real-time processes such as flood monitoring. The NRFA has slightly different requirements, with a need for tools to summarise, visualise and analyse daily mean flows rather than converting raw water level data to flows via ratings, and so has traditionally developed its own bespoke systems. This approach allows for flexibility in the development of new aspects of the system, particularly for the production of additional statistics and the addition of new data validation and improvement tools. The NRFA systems also allow greater integration of time series with spatial information than many existing tools. The systems consist of databases, computer code for data handling and statistical calculations, software for visualisation within user interfaces, and external web-pages for the dissemination of data. An incremental approach to development is taken, whereby new functionality identified for addition is built based on common core functionality, allowing the reuse of standard building blocks of existing code. The development and maintenance of bespoke systems requires the retention of specialist skills in software development and databasing, and so is not an option for many NHSs. However the basic functionality outlined below should be of interest to anyone involved in the management of hydrological archives.

Hydrological data management within the NRFA, particularly the annual on-take, validation and storage of river flow data, is undertaken within a set of specific guidelines and under a Service Level Agreement with the Measuring Authorities through which data quality and quantity submitted is monitored. The NRFA systems are integrated with these guidelines and provide the means for achieving the high levels of validation required.

Time series datasets

The principal datasets within the NRFA are gauged daily mean flow series from gauging stations across the U.K. For several stations, series of estimated naturalised flow series are also archived. Some longer records include periods of monthly mean flows where daily data are not available. Currently the archive constitutes over 50 000 station years of monthly and daily mean flow data for some 1400 gauging stations. Monthly peak flows are also archived. Hydrological context to these river flows, to enable validation and other NRFA activities, is provided through additional datasets. These include daily and monthly storage raingauge data from the UK Met Office, groundwater levels from the British Geological Survey, reservoir storages, etc. Further derived time series such as monthly catchment average rainfall and runoff are either pre-processed and stored, or derived at the point of use.

Metadata

Additional datasets, which could be described as metadata, provide information on gauging stations and gauging station catchments. These generally include anything that could assist users of the river flow data, within or external to the NRFA, to understand the data, their accuracy and major influencing factors. Gauging station information includes grid references, operator information, gauging station structure, bankfull flow, sensitivity, etc. Time series metadata, much of which is automatically derived, includes completeness information and key time series statistics (mean flow, exceedance percentile flows, Base Flow Index, etc.). A station description provides a digested summary of known influences on the flows such as weed growth, bypassing, major abstractions, discharges or reservoirs, and aspects of the station's history, including significant rating changes, alterations to measurement
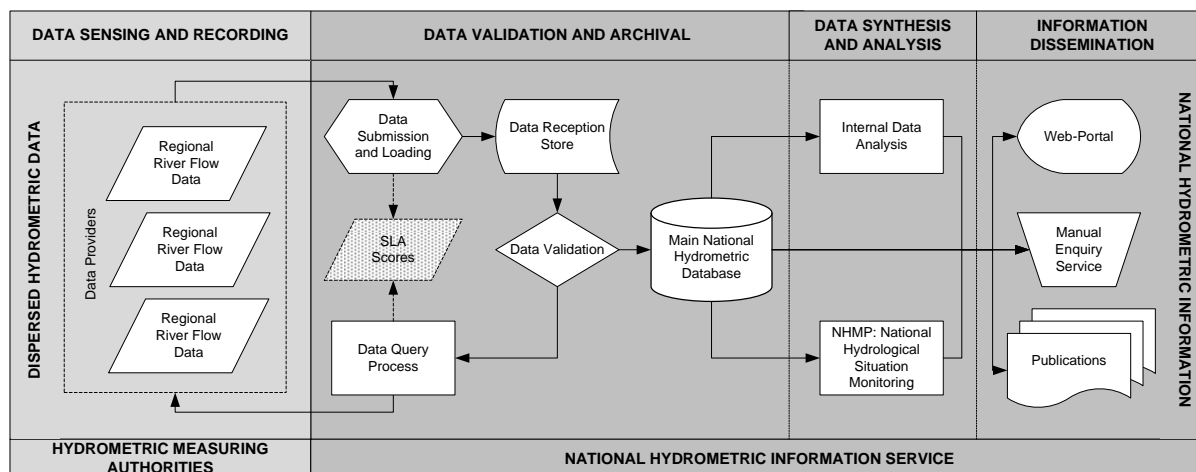

Figure 1   *Schematic representation of NRFA data validation, analysis and dissemination processes*
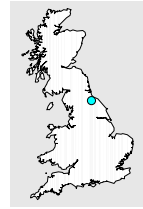
**NRFA Gauging Station Information Sheet**

Leven at Leven Bridge

Measuring Authority: Environment Agency
Grid Reference: 45 (NZ) 445 122
Station Type: Compound broad-crested weir

Gauged Flows and Rainfall: 1959-2008
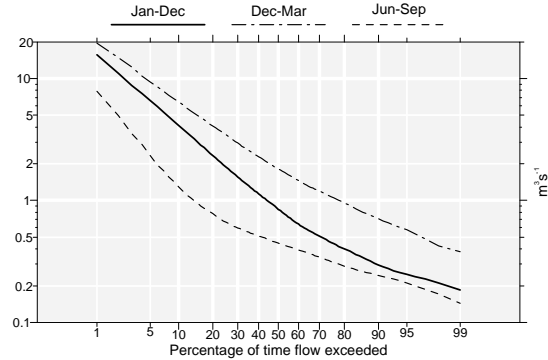NRFA Station Number: 25005
Local Number: 3802

Daily Flow Hydrograph

Max. and min. daily mean flows from 1959 to 2008 excluding those for the featured year (2007; mean flow: 1.98 m³s⁻¹)
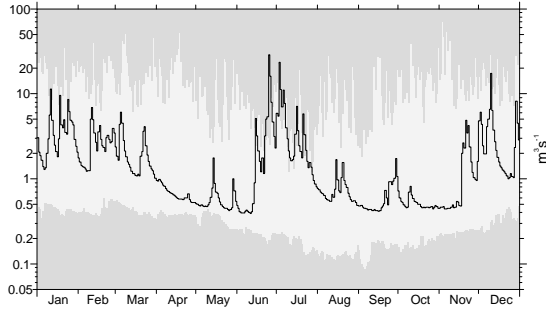
Flow Duration Curve

Flow Statistics

(Units: m³s⁻¹ unless otherwise stated)

Mean flow	1.88	
Mean flow (ls⁻¹/km²)	9.58	
Mean flow (10⁶m³/yr)	59.3	
Peak flow / date	124.6	2 Nov 2000
Highest daily mean / date	71.6	2 Nov 1968
Lowest daily mean / date	0.088	4 Sep 1976
10 day minimum / end date	0.106	9 Sep 1976
60 day minimum / end date	0.146	9 Sep 1976
240 day minimum / end date	0.323	6 Jan 1965
10% exceedance (Q10)	4.119	
50% exceedance (Q50)	0.846	
95% exceedance (Q95)	0.248	
Mean annual flood	45.4	
IH Baseflow index	0.44	

Rainfall and Runoff

	Rainfall (1959-2008) mm						Runoff (1959-2008) mm					
	Mean	Max/Yr		Min/Yr			Mean	Max/Yr		Min/Yr		
Jan	64	132	1960	13	1989		48	99	2008	7	1989	
Feb	50	104	1978	6	1998		36	101	1978	9	1973	
Mar	54	159	1999	6	1973		36	142	1979	6	1973	
Apr	56	179	2000	10	2007		29	113	2000	5	1997	
May	52	151	1979	14	1992		17	76	1979	6	1982	
Jun	62	164	2007	13	1960		12	44	2007	3	1960	
Jul	62	152	1960	16	1983		10	57	2007	3	1976	
Aug	72	155	2004	12	1995		11	51	1978	2	1976	
Sep	64	251	1976	14	1997		11	63	1993	2	1990	
Oct	66	185	1960	11	1969		20	112	1960	3	1959	
Nov	75	181	2000	19	1964		31	157	2000	3	1964	
Dec	70	156	1978	21	1988		41	92	1979	6	1964	
Year	747	998	2000	448	1964		302	542	2000	94	1989

Station and Catchment Characteristics

Station level	(mOD)	5.2
Sensitivity	(%)	23.0
Structurefull flow		70.00
Catchment area	(km²)	196.3
Maximum altitude	(mOD)	454
FSR slope (S1085)	(m/km)	4.17
1961-90 rainfall (SAAR)	(mm)	725
FSR stream frequency (STMFRQ)	(junctions/km²)	
Urban extent	(0-1)	0.0099
Flood Attenuation Index	(0-1)	0.9977

Station and Catchment Description

Compound broad-crested weir, width 17.4m, with a bypass Crump profile weir width 4.6m. Theoretical rating (but also gauged to almost Qmed). Sharp bend and road bridge just u/s of weirs and large drop below, therefore believed to be modular throughout range. Suffers from siltation upstream of the weir which may affect calculated flows. Bypassed only during Nov. 2000 flood (peak flow may be unreliable). Oct 2005 data removed due to construction of fish pass. Flows from 04/08/2005 to 06/01/2006 may also be affected, and were due for investigation in spring 2007. Sensibly natural regime. Due for replacement with new station at Foxton Bridge, 6km upstream, commissioned in 2005.

Headwaters drain from the Cleveland Hills. Mixed geology of mostly Permian/Jurassic age, generally low permeability; 80% superficial deposits, mainly boulder clay. Mixed agriculture, with some moorland and forestry in headwaters, minor urban development in

Factors Affecting Runoff

- Natural to within 10% at the 95 percentile flow.
- Runoff increased by effluent returns.

Summary of Archived Data

Gauged Flows and Rainfall

	0 1 2 3 4	5 6 7 8 9
1950s	- - - - -	- - - - e
1960s	A A A A A	A A A A A
1970s	A A A A B	A A E A A
1980s	A A A A A	A A A A A
1990s	A A A A A	A A A A A
2000s	A D A A A	D A A E

Key:

	All rainfall	Some or no rainfall
All daily, all peaks	A	a
All daily, some peaks	B	b
All daily, no peaks	C	c
Some daily, all peaks	D	d
Some daily, some peaks	E	e
Some daily, no peaks	F	f
No gauged flow data	=	-

Naturalised Flows

Key:

All daily, all monthly	A
Some daily, all monthly	B
Some daily, some monthly	C
Some daily, no monthly	D
No daily, all monthly	E
No daily, some monthly	F
No naturalised flow data	=

Figure 2   *NRFA gauging station summary sheet showing the range of metadata provided to data users*

structures, or periods with known data quality issues, considered by the NRFA to be most important for end users. Catchment information includes statistics on rainfall, land cover, elevation, hydro-geology, and a range of statistics obtained from the Flood Estimation Handbook (Bayliss, 1999), and are derived from the spatial datasets described below. Figure 2 shows a standard NRFA gauging station information sheet illustrating the range of gauging station metadata available.

These metadata provide the context for expert and novice users of hydrological data to gain an initial understanding of the quality of flow data from a gauging station, and the maintenance of these metadata are key to the NRFA's efforts to maximise the utility of gauging station data. To this end, selected statistics and metadata are published on a regular basis within a Hydrometric Register (Marsh and Hannaford, 2008) and provided to data users.

Spatial datasets

River flows are the result of hydrological processes distributed across the area of a catchment and can provide valuable information about those processes. Conversely, to understand the river flow regime, information about the catchment is required, particularly aspects of the catchment important to the hydrological processes. The statistical catchment metadata described above can assist an understanding of a catchment's influence on flows, but spatially distributed data provide a far more comprehensive view. The NRFA maintains several core spatial datasets, specifically, gauging station locations and their digitised catchment boundaries, and further sets of contextual data that allow the river flow data to be fully understood. These additional datasets include land cover (Fuller *et al.*, 2002), hydrogeology, a digital terrain model (DTM) (Morris and Flavin, 1990), digital river network, Standard Average Areal Rainfall (SAAR) (Spackman, 1993), and a range of the FEH Catchment Descriptor datasets, many of which are indicative statistics derived from the DTM and Hydrology of Soil Types (Boorman *et al.*, 1995) datasets.

NRFA databases

Databases within a hydrological archive must provide some fundamental functionality: security, auditing, and flexibility of access. To this end, NRFA data are stored in an Oracle Relational Database Management System, backed-up daily. The database design provides a simple and transparent model allowing fast and reliable access via direct query or through software tools. This model contrasts with those of many commercial hydrological data management tools, which tend to use bespoke data formats or very complex RDBMS structures, and allows application code to be rapidly developed to utilise the data directly, including libraries of SQL commands for specific data retrievals, R scripts for direct statistical analysis of the data, Fortran routines managed by hydrological modellers, and complex software tools. This enhances the utility of the river flow data within a research organisation such as CEH.

Time series data, provided by the Measuring Authorities, are loaded to temporary tables during the validation process. Where available, data quality flags and comments from Measuring Authorities are also loaded to assist the validation process with information recorded at the point of measurement concerning issues occurring with equipment, maintenance of stations, etc. This allows NRFA staff validating the data to understand data issues

without recourse to MA staff in the first instance. During the validation process, software tools allow more problematic queries about specific flow series to be passed to the MA staff for explanation or resolution. Validated data are loaded from the temporary tables onto the main database and merged with the archive. Auditing of changes to data within the database is enforced, ensuring a minimal audit trail of all changes to the data. Spatial data are mostly managed within an Oracle database using the ESRI Spatial Database Engine (ESRI 2010), providing secure storage, multi-user access, and version management.

Time series data management and analysis

Perhaps the most important aspect of the NRFA data management systems is the management and analysis of time series data. The assessment of hydrological conditions and events and placement of these within the long-term hydrological context requires the processing and review of huge ranges of statistics across different intervals. Fast and accessible tools for the manipulation of time series data underpin much of the NRFA work in this area, particularly for publications such as the National Hydrological Monitoring Programme's monthly Hydrological Summary (http://www. ceh.ac.uk/data/nrfa/water_watch.html), and occasional reports such as those on the 2007 floods (Marsh and Hannaford, 2007) and the 2004 – 2006 drought (Marsh, 2007), and papers on trends in river flow statistics (e.g. Hannaford and Marsh, 2008). The code for these tools forms the core of the NRFA time series systems. Examples of the functionality provided include conversion between intervals, for example daily flows to monthly or annual mean, maximum, minimum, or total flows, calculation of *n*-day statistics, application of mathematical operations across multiple time series, as well as more specific hydrological functions such as base flow calculation and separation, and flow to runoff conversion. The code is written in such a way as to allow the output of one function to form the input of another, meaning that a complex chain of conversions and operations can be defined by the user to produce a huge range of statistics without the need for predefined functions.

A further, and perhaps more important, aspect of the code design, is the generic approach to time series handling. The code is based around core concepts of time series management, such as data type (including flow, rainfall, etc.), unit (m^3s^{-1}, mm, etc.), interval (15-minute, daily, monthly, etc.), statistic (mean, max, min, total, etc.). This allows new time series datasets to be simply integrated into the NRFA systems and analysed with the same tools. In addition, new datasets do not need to conform to a specific database format or structure, as long as the meaning of these concepts can be specified in relation to each time series. Using tools initially only available for river flows and rainfall, CEH staff can integrate NRFA data with existing databases of groundwater levels, water quality data, and lake levels, as well as project specific databases containing model outputs or other time series data, without the need for restructuring the data. This integration not only allows the tools to be applied to new datasets, but enhances the ability to validate river flows through the provision of more information on the hydrological context to the flows. The UK Measuring Authorities can currently produce time series data within a standardised xml file format which uses the same concepts and the NRFA can automatically ingest and analyse any data within this format. NRFA flow data have recently been integrated with the UK Harmonised Monitoring water quality database (Simpson, 1978), and national-level assessment of, for example, chemical mass loads can be rapidly automated.
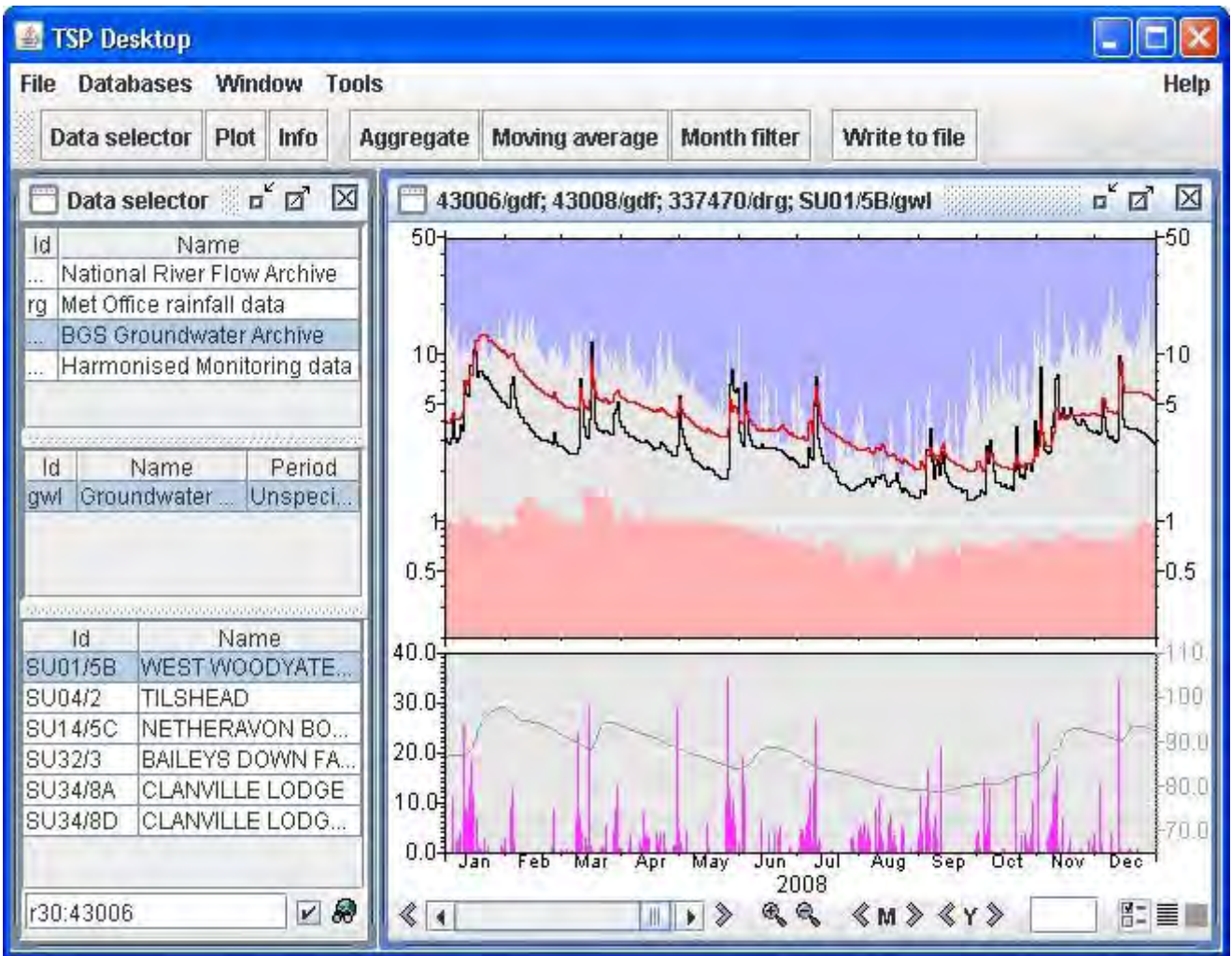
Figure 3   *Time Series Plotter showing graph of river flows from the river Nadder (black line) and historic daily extremes for this series (blue and red banding), together with flows from the river Wylye (red line), rainfall from a raingauge within the catchment (pink bars), and a nearby groundwater level record (grey line).*

Figure 4   *Time Series Plotter showing graphs and tables of mean annual flow (red line) and annual 7-day minimum flows (black points) for the Nadder at Wilton, against total annual catchment annual rainfall (green bars). Table shadings indicate the highest (blue), and lowest (red), 10% of values.*

An important recent addition is a set of tools for infilling missing data. Missing data seriously undermine the utility of river flow data for purposes such as statistical analysis and model calibration. For example, many methods for the extraction of annual maxima or minima will reject years with missing values, regardless of the likelihood of the missing periods containing extreme flows. To manually assess this likelihood when producing statistics is time-consuming and requires access to other contemporary information. Appropriate and informed infilling of missing data, in particular with an aim to maximising the number of complete years data, can dramatically improve the utility of a time series for many purposes. Guidelines for the application of various infilling methods are being developed within the NRFA (Harvey *et al.*, 2010) and included as tools within the NRFA systems.

Time series data visualisation

The validation of river flow data is a difficult process to automate. Basic errors within data can be identified through automated procedures but the natural variability, and the subtlety of many hydrological data issues, means that a visual review of data is an essential aspect of comprehensive secondary validation. The NRFA validates each new year of flow data through the comparison of key statistics for the new period with those for the entire period of record, and a visual comparison with earlier records at the same location, with data from nearby locations, and with other hydrological data such as rainfall. The tool developed for visualisation, Time Series Plotter (TSP) provides the functionality for this. Based on the time series management code described above, TSP allows the integration of time series data and statistics in a flexible 'drag-and-drop' user interface. Figure 3 shows TSP displaying one year of river flow data against other nearby river flows, rainfall and groundwater levels. Figure 4 shows TSP displaying graphs of longer-term statistical information from this station, together with a table of the same information. The means for rapidly producing such plots and statistics provides a powerful capability to identify data issues within individual events and long-term series. The TSP tool allows NRFA staff validating flow series to rapidly derive new series, using tools for calculating baseflow, *n*-day means, etc. and display these alongside raw data. Time series generated from infilling methods can be viewed alongside the series to be infilled and flows from nearby catchments to assess the validity of the generated series prior to application. The flexibility and ease with which raw data and derived series can be produced and integrated allows users to obtain a full understanding of events.

Metadata, including gauging station quality criteria and catchment information, can also readily be accessed through the system, providing further information to enable decisions to be made regarding the quality of river flow data.

Spatial data visualisation and analysis

Spatial data and spatial information are also heavily used within NRFA processes. Visually this is important for the assessment of the spatial variability of aspects such as elevation, land cover, rainfall within catchments, as well as other aspects that can affect runoff processes and influence river flow data. Visualisation is generally performed using ESRI's ArcGIS, which allows for the rapid integration and querying of all of the NRFA spatial datasets available. Spatial analysis is performed in a number of ways. Individual queries are run within the GIS software to undertake a specific piece of analysis. Where analyses are regularly run, or are run for many locations or catchments, these can be scripted and built into tools. Several series of complementary tools exist for performing common exercises such as catchment boundary definitions, retrieval of catchment statistics, etc. The potential for executing analyses across thousands of catchments allows publications such as the UK Hydrometric Register to be produced, and can provide information to feed into modelling studies.

A subset of NRFA spatial analysis tools provides information for the assessment of gauging station networks. The Catchment Utility Index (CUI) (Laize *et al.*, 2008) provides an indication of the frequency with which a gauged catchment in the UK is called into the Flood Estimation Handbook pooling group for 4 million catchments across the country. This can provide an independent measure of the strategic usefulness of gauging stations, though focused on floods and using one particular, albeit national standard, method. The CUI tools that produce this index are an example of the NRFA's capability to provide an independent national perspective on the gauging station network. A second subset of spatial analysis tools provides information for assessing how well a given gauging station catchment represents a region as a whole. Laize (2004) describes the Representative Catchment Index (RCI), which can be calculated for one or more given catchment descriptors, to establish 'representativeness' or a target region in terms of combinations of e.g. land use, elevation and geology. The output of the RCI tools is an independent measure of the utility of a catchment within a region, or nationally, and can be used within the strategic management or development of gauging station networks, for example identifying valuable additional gauging station locations.

These catchment analysis tools are complex, but produce simple and readily understandable indices that provide important national capability within the UK. Adapted to local requirements and datasets, they would be useful for NHSs in other countries for the assessment and analysis of gauging station networks.

Dissemination of river flow data

The dissemination of river flow data forms an important part of the mandate of the NRFA, and is an important aspect of the work of hydrological archives in other countries. Traditionally, this dissemination has taken the form of manual requests and retrievals. More recently, web-based data access has become popular and is now the standard means for access to data. The new NRFA website aims to provide the users of UK river flow data with not only data, but also the full range of information required to understand the river flows, their origins, and the appropriateness for application to a given purpose. Web-pages for each gauging station provide NRFA metadata, as described above, allowing users to assess the quality and extent of the flow record prior to downloading. For many stations dynamic graphing is available for river flow time series and flow duration data. Spatial data pages provide maps and statistics on the range of catchment descriptors available. Figure 5 shows some example pages for one station. A map-based portal allows users to find data by geographic location, or by searches based on gauging station, catchment or time series metadata, for example identifying stations with mean flow greater than 50 m^{3}s^{-1}, or where the catchment land cover is >50% forest. The time series graphing and search pages use extensions of the core time series management and visualisation code used in the desktop tools to access dynamically the NRFA databases. This approach allows new core functionality to be rapidly

Figure 5   *Examples of gauging station data pages from the NRFA website.*

delivered via the internet, and ensures publicly available data is always up-to-date.

Summary and future developments

The UK National River Flow Archive maintains a wide variety of datasets and hydrological data management systems in order to undertake its work in maximising the utility of UK river flow data and providing commentary on UK hydrological conditions and gauging station network development. The principal power of these systems lies in their ability to rapidly integrate a wide variety of time series datasets within a flexible visual interface and analysis tools. Furthermore, comprehensive spatial datasets provide information on gauging station catchments, essential for the understanding of river flow series, and can feed into tools for assessing the utility and representativeness of catchments within wider, e.g. national, networks.

Future plans are to improve metadata management, particularly providing a greater range of expert comments on river flow time series and individual events, and to broaden the suite of spatial data analysis tools to include river network analysis to allow the linking of river-based sites from different networks, and furthering the gauging station network analysis tools.

The systems described above, whilst specific to the situation and requirements of the UK, can be seen to provide a model for data management systems that would complement national hydrological archives around the world. Tools for the integration of time series datasets, the production of statistics, and the visualisation of river flow time series alongside other datasets describing hydrological conditions, are essential for the detailed secondary validation required for improving the utility of river flow data. The availability of spatial datasets such as Digital Terrain Models for catchment definition and understanding of topographic variability, land cover, geology, etc. is also essential to adequately understand hydrological processes influencing river flows. Tools for analysing these spatial datasets to produce the type of statistics required for understanding of gauging station networks as a whole could improve the management of networks in many countries.

References

Bayliss, A.C. 1999. Catchment descriptors. *Volume 5 of the Flood Estimation Handbook*. Centre for Ecology, Wallingford.

Boorman, D.B., Hollis J.M. and Lilly A. 1995. Hydrology of Soil Types: A Hydrologically Based Classification of the Soils of the United Kingdom. Institute of Hydrology, *Report No. 126*.

CEH. 2010. HYDATA software web page. http://www.ceh.ac.uk/products/software/CEHSoftware-HYDATA.htm. Accessed April 2010.

ESRI. 2010. ArcSDE software web page. http://www.esri.com/software/arcgis/arcsde/index.html. Accessed April 2010.

Fuller, R.M., Smith, G.M., Sanderson, J.M., Hill, R.A., and Thomson, A.G. 2002. The UK Land Cover Map 2000: Construction of a Parcel-Based Vector Map from Satellite Images. *The Cartographic Journal*, **39**, 15–25.

Harvey, C.L., Dixon, H., and Hannaford, J. 2010. Developing best practice for infilling daily river flow data. *This volume*.

Isodaq. 2010. HydroLog software web page. http://www.isodaq.co.uk/datasystems/hydrolog4. Accessed April 2010.

Kisters. 2010a. WISKI software web page. http://www.kisters.net/english/html/Nav_NA/RMS/WaterManagement.html&center=8FB26FB39796717FC125737F005CE7CB. Accessed April 2010.

Kisters. 2010b. Hydstra software web page. http://www.kisters.net/english/html/Nav_NA/RMS/WaterManagement.html&center=8835FC6A8D9A9A0DC125732F003F36CC. Accessed April 2010.

Laize, C.L.R. 2004. Integration of spatial datasets to support the review of hydrometric networks and the identification of representative catchment. *Hydrol. Earth Syst. Sci.*, **8**, 1103–1117.

Laize, C.L.R., Marsh, T.J., and Morris, D.G. 2008. Catchment descriptors to optimise hydrometric networks. *Proc. Instn. Civil Engrs, Water Management*. **161**, 117–125.

Marsh, T.J. 2002. Capitalising on river flow data to meet changing national needs — a UK perspective. *Flow Measurement and Instrumentation*. **13**, 291–298.

Marsh T.J. 2007. The 2004-06 drought in southern Britain. *Weather*. **62**, 191–196.

Marsh, T.J. and Hannaford, J. 2007. The summer 2007 floods in England and Wales. *National Hydrological Monitoring Programme series*. CEH/BGS. 30 pp.

Marsh, T.J. and Hannaford, J. (Eds). 2008. *UK Hydrometric Register*. Centre for Ecology and Hydrology/British Geological Survey. 216 pp.

Morris D.G. and Flavin R.W. 1990. A digital terrain model for hydrology. *Proc. 4th Int.Symp. on Spatial Data Handling*, *Zurich*, **1**, 250–262.

NIWA. 2010. Tideda software package. http://www.niwa.co.nz/our-services/software/tideda. Accessed April 2010.

Simpson, E.A. 1978. The Harmonization of the Monitoring of the Quality of Inland Fresh Water. *J. Instn Water Engrs Sci.*, **32**, 45–56.

Spackman, E. 1993 Calculation and mapping of rainfall average for 1961–1990. *Proc. British Hydrological Society Symp*, Salford.