Hydrology Research | 43.5 | 2012

An appraisal of the performance of data-infilling methods for application to daily mean river flow records in the UK

Catherine L. Harvey, Harry Dixon and Jamie Hannaford

ABSTRACT

River flow records are fundamental for the sustainable management of water resources and even very short gaps can severely compromise their utility. Suitably-flagged flow estimates, derived via judicious infilling, are potentially highly beneficial to data users. The UK National River Flow Archive provides stewardship of, and access to, UK river flow records. While many datasets held on the archive are complete, gaps remain across a wide range of flow records. A comprehensive assessment of existing techniques for infilling these gaps is currently lacking. This paper therefore assesses 15 simple infilling techniques (including regression, scaling and equipercentile approaches), each relying upon data transfer from hydrologically-similar donor stations, to generate estimates of flow at 26 representative gauging stations. Results reveal the overall superiority of equipercentile and multiple regression techniques compared to the poorer capability of catchment area scaling. Donor station choice has a strong influence on technique performance. Modifying datasets to improve homogeneity, by seasonally grouping flows or excluding certain periods, offers improved performance. These findings provide a foundation upon which guidance on infilling river flow records can be based in future, allowing hydrometric practitioners and data end-users alike to adopt a consistent and auditable approach towards infilling.

Key words | equipercentile, hydrometric data, infilling, missing data, river flows, time series

Catherine L. Harvey
Harry Dixon
Jamie Hannaford (corresponding author)
Centre for Ecology and Hydrology,
Maclean Building,
Crowmarsh Gifford,
Wallingford,
Oxfordshire,
OX10 8BB,
UK
E-mail: *jaha@ceh.ac.uk*

INTRODUCTION

River flow records are a vitally important asset and their completeness forms a crucial aspect of their utility. Even very short data gaps can preclude the meaningful calculation of important summary statistics and hydrological indicators, such as monthly runoff totals or *n*-day minimum flows, thus inhibiting the analysis and interpretation of past flow variability. River flow records are also a vital input to hydrological models, including those used for predicting future behaviour (Hannah *et al.* 2011); gaps can have a deleterious impact on estimates derived from prediction and forecasting tools. Complete records are therefore critical to the sustainable management of water resources worldwide, and gaps in records represent a loss of information which can potentially affect the interpretation of data, and the scientific outcomes of analysis; Marsh (2002) argues that, in many cases, the inclusion of suitably flagged flow estimates is preferable to leaving gaps in records.

Within the UK, the National River Flow Archive (NRFA) acts as the main hydrometric archive, collating data from different monitoring network operators. Daily mean river flows are stored for over 1,500 gauging stations and validated, analysed and disseminated to a wide range of users (Dixon 2010). While the majority of these flow records have high overall percentage completeness (78% of stations have records that are at least 95% complete; Marsh & Hannaford 2008), closer inspection reveals a significant quantity of both contemporary and historical gaps, ranging in length from a single day to several months. Such gaps are an inevitable consequence of factors such as essential gauging station maintenance, equipment malfunction, changes in instrumentation, data processing

doi: 10.2166/nh.2012.110

issues and human error. For some gaps, the data are likely to be unrecoverable; for example, an extreme high flow event that destroys a gauging station may be difficult to estimate with any degree of certainty. In most cases, however, gaps may be amenable to infilling, particularly where hydrological conditions are relatively stable.

A previously observed decline in the completeness of river flow data submitted to the NRFA (Marsh 2002) can in part be attributed to a lack of standardised infilling guidance which, in its absence, has discouraged the infilling of gaps. While there has been a demonstrable improvement in completeness in recent years (Dixon 2010), historical data gaps remain and short sequences of missing daily mean flows (which appear readily amenable to infilling) still regularly occur in data submitted to the NRFA. This highlights a need for informed guidance on the use of infilling techniques to promote a consistent, repeatable approach towards such record gaps. Simple, quick-to-apply techniques that perform well across an extensive range of catchments could find wide applicability, thus limiting the investment of time and resources required to infill data to an appropriate degree of accuracy, while also significantly enhancing the overall utility of time series. However, there are currently no widely-accepted standard techniques for data infilling, either in the UK or internationally.

The aim of this paper is to evaluate the performance of a range of existing simple methodologies for gap filling. A variety of catchment types in the UK are used, in order to test the applicability of such techniques across a broad spectrum of hydrological settings. This testing framework is crucial as the aim is not to find a 'one size fits all' methodology; rather to assess the range of applicability of the multiple techniques and their limitations. Similarly, it is recognised that infilling will not be appropriate in many situations, so the aim is to find mechanisms which show good general applicability over the flow range, rather than to identify specific instances where such methods could and should be applied. The overall aim is to identify infilling mechanisms which demonstrate accuracy and versatility, for future application alongside expert judgment.

The appraisal presented in this paper is an important first step in the development of guidance on data infilling for hydrometric measuring authorities. It is anticipated that this approach will also hold relevance for the wider hydrometric data user community both within the UK and internationally, and may feed into future developments in international protocols for data management (e.g. World Meteorological Organization 2008). Systematic reviews of data-infilling techniques are rare, and (to the authors' knowledge) no previous study embraces such a range of techniques (15) across such a number of cases (26 catchments UK-wide).

The paper is structured as follows. Firstly, a review of published techniques is presented. The UK river flow data and the methodology used to quantify the performance of existing techniques are then described. A results section follows, drawing out the key findings from an intercomparison of all techniques. The applicability of these techniques in practice is then demonstrated, firstly by reference to examples that illustrate particular issues with practical application of the methods, and then through two case studies of infilling applied to catchments which were not in the original dataset used for technique appraisal.

REVIEW OF EXISTING INFILLING TECHNIQUES AND STUDIES

Many papers in the literature undertake some form of river flow data infilling, which is often done rather casually, without adequately describing how the infilling was completed or assessing its effects on the results. However, there are also a number of specific methodologies for infilling which have been advocated in the literature. Existing techniques, developed either exclusively for infilling or alternatively for flow record extension, were assessed for appropriateness (Table 1). Most techniques rely upon functions for transferring data from other gauging stations. Such stations are referred to as 'donors', whilst the term 'target' indicates the station record that requires infilling. The qualities that constitute a useful donor are arguably a research topic in their own right; the topic of 'regionalisation' of hydrological variables, i.e. their extrapolation in space, is a fertile area of research with a long history. There are a wide range of approaches to donor selection even in the UK (for general reviews, see Wagener *et al.* 2007; Shaw *et al.* 2010), with no single technique appropriate for the full range of flows. Common considerations include proximity and similarity

Table 1 | Summary of common infilling techniques

Method	Summary	References
Manual inference	Estimates are derived through visual comparison with donor flows. Accuracy should be fairly assured for short gaps with no rainfall or longer gaps during stable recessions, but other conditions may lead to increased difficulty and subjectivity in determining estimates.	Rees (2008)
Serial interpolation techniques	These include linear, polynomial or spline interpolation and are likely to only be successful throughout stable periods.	Rees (2008)
Scaling factors	Donor flows are multiplied by a scaling factor, such as the ratio of the target and donor catchment areas or a weighting based upon the distance between the target and donor.	Kottegoda & Elgy (1977); Wallis *et al.* (1991); Hughes & Smakhtin (1996)
Equipercentile technique	The donor and target flow percentile values are assumed equal for any given day. Gaps are infilled by calculating the donor flow percentile values and using the existing target flow data to derive the target flows equivalent to these percentile values.	Hughes & Smakhtin (1996); Smakhtin & Masse (2000); Rees (2008)
Linear regression	A regression equation between the target and at least one donor is derived, commonly via the least squares method, and used to calculate absent target flows. Flows may first be transformed, for example, logarithmically.	Raman *et al.* (1995); Hirsch (1979, 1982)
Hydrological modelling	This varies from black-box modelling, whereby the model inputs are related to the outputs without considering the processes involved, to the more complex process-based models and use of artificial neural networks.	Khalil *et al.* (2001); Ilunga & Stephenson (2005); Beven (2012)

to the target catchment in terms of hydrological responsiveness, climate and catchment physiography (Rees 2008). Where available, multiple donors can enhance the likelihood of capturing the many influences impacting a target's flow regime, but a single donor could be sufficient if it has a similar hydrological regime, which is more likely if located very close to the target or on a major upstream tributary (Hughes & Smakhtin 1996).

In relation to the issue of donor selection, there is much scientific debate in the literature on regionalisation as to which mechanisms (and which catchment attributes) should be used to index catchment similarity (e.g. McIntyre *et al.* 2005; Yadav *et al.* 2007) or even whether to use catchment similarity measures as opposed to local data transfer (Merz & Blöschl 2005). However, there is currently no agreed framework for catchment similarity classification in hydrology (Wagener *et al.* 2007), and the concept of 'uniqueness of place' (Beven 2000) – whereby catchments are unique in terms of their topography, soils, rock types, vegetation and anthropogenic modification – arguably limits the potential for such generalisation. The present study will therefore not address donor selection criteria, but will attempt to test infilling mechanisms on as wide a range of donor/target pairs as possible, to determine which methods work best across a range of situations.

Earlier studies of infilling techniques have focused on either a single technique or a small number of techniques all belonging to the same general approach (for instance, regression techniques: Hirsch 1982; scaling techniques: Kottegoda & Elgy 1977). A novel aspect of the present study is the consideration of a large number of techniques, encompassing a broad range of possible approaches.

To date, the majority of studies have limited their analyses to a small number of case-study targets (for example: Gyau-Boakye & Schultz 1994; Elshorbagy *et al.* 2000; Amisigo & van de Giesen 2005). Hughes & Smakhtin (1996) considered a larger sample, but only a single infilling technique was tested. Across the UK, the marked variability in hydrological regimes and the prevalence of anthropogenic influences (Marsh 2002) necessitates consideration of a high number of target stations, in order to reliably determine whether a technique is widely applicable. This study will therefore test techniques on a sample of 26 hydrologically representative UK gauging stations.

The relative performance of infilling techniques can be compared through infilling artificially created gaps (for example: Gyau-Boakye & Schultz 1994) but, despite careful selection to reflect diverse conditions, this methodology is still dependent upon the nature and magnitude of the time series when the gaps are established. An alternative approach, followed by the present study, is to compare the ability of techniques to simulate entire target flow records (for example, Elshorbagy *et al.* 2000), thus indicating which techniques can be expected to perform better for any given gap, across the flow range.

METHODOLOGY AND DATA

There are three factors which are likely to influence the reliability of data infilling: (1) the nature of the donor station(s), (2) the location and duration of the gap, and (3) the infilling procedure. As the aim of this study is to compare a wide range of infilling mechanisms (item 3), this study seeks to control (1) and (2) insofar as possible, by selecting a wide range of donor–target situations, and testing the infilling methods across the whole flow range rather than for particular gaps. An intercomparison is therefore made of 15 different techniques, for 26 donor–target combinations. For each technique, a full daily mean time series was simulated using the observed donor flow time series. The utility of techniques is assessed using three indicators of performance. The following sections describe the process in detail.

Infilling techniques

This study assesses the utility of 15 infilling techniques, including equipercentile, scaling and regression approaches, all of which exploit data transfer from either one (single donor techniques) or two (dual donor techniques) other gauging stations (Table 2). Prior to applying infilling techniques, datasets can be modified to potentially improve technique performance. For example, separating time series into monthly or seasonal divisions can result in more homogeneous flow groups (for example, Raman *et al.* 1995),

Table 2 | Infilling techniques tested by this study. In order to account for any flow records containing zero flows, the log-transformation took the form of ln(flow+1). For seasonal flow groupings, December–February flows were grouped for Winter, March–May for Spring, June–August for Summer and September–November for Autumn. Techniques were applied to datasets comprising days when observed flows existed for both the target and primary donor (single donor techniques) or all three stations (dual donor techniques)

Acronym	Name	Details
LR	Linear regression	Least-squares linear regression between target and primary donor flows.
LR Seas	Linear regression seasonal	As above but using seasonally grouped flows.
LR Log	Linear regression log	Least-squares linear regression between log-transformed target and primary donor flows.
LR LS	Linear regression log seasonal	As above but using seasonally grouped flows.
M1	MOVE.1	MOVE.1 regression between target and primary donor flows (Hirsch 1982).
M1 Log	MOVE.1 log	As above but using log-transformed flows.
Equi	Equipercentile	Equipercentile technique applied using primary donor flow percentile values.
CA	Catchment area scaling	Catchment area scaling applied using target and primary donor catchment areas.
LTM	Long-term mean scaling	Long-term mean scaling applied using target and primary donor long-term mean flow values.
LTM Seas	Long-term mean scaling seasonal	As above but using seasonal groupings of flows.
MR	Multiple regression	Least-squares linear regression between flows of target and both donors.
MR Seas	Multiple regression seasonal	As above but using seasonally grouped flows.
MR Log	Multiple regression log	Least-squares linear regression between log-transformed flows of target and both donors.
MR LS	Multiple regression log seasonal	As above but using seasonally grouped flows.
W.Equi	Weighted equipercentile	Equipercentile technique applied using each of the donors and averaging the resulting estimates for each date.

which may also address the common non-stationarity of flow records (Hirsch 1979). Other data preconditioning can include the application of a log-transformation to flow series to reduce skewness in the distribution of the data (for example, Hirsch 1979, 1982). The chosen techniques therefore feature variations of the same approach based upon first log-transforming and/or seasonally grouping flows, with the latter reflecting a compromise between the reduced sample size effected by grouping data, the potential improvement such grouping could afford and the computational demands of seasonal versus monthly grouping.

Despite its potential to offer highly accurate estimates, hydrological modelling was not considered, since such methods are too resource-intensive for rapid application to a large number of stations; the results would be very dependent on the choice of model used, limiting their utility for developing generic guidelines in future. Simple manual inference and serial interpolation techniques were also ignored as, despite their undoubted practical utility in certain circumstances, especially short gaps, they are heavily reliant upon subjective decisions and cannot be easily automated and objectively compared within the testing framework used in this study. A final criterion was to utilise only river flow data sources in the infilling process, avoiding dependence upon other datasets (in particular, catchment rainfall) which may not always be readily available to users.

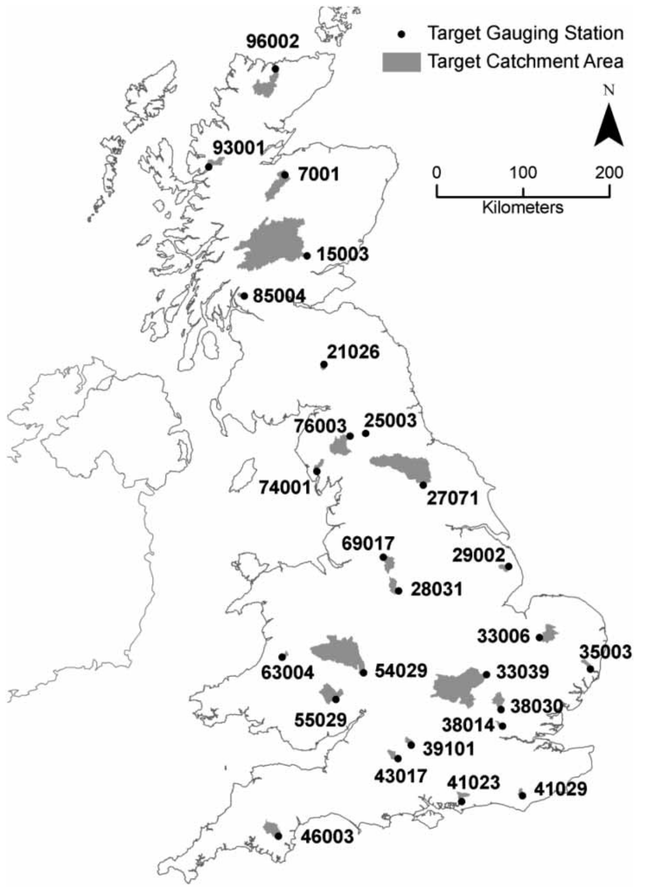

Figure 1 | Target gauging station locations, depicted by catchment area and labelled according to NRFA station ID.

Intercomparison dataset

The 26 target stations were selected from the NRFA to provide a broad spatial distribution across the UK and incorporate both very responsive and baseflow-dominated, large and small, and natural and artificially influenced catchments (Figure 1). For each target, a primary donor station was selected for use with the single donor techniques and an additional secondary donor station for use with the dual donor techniques (Table 3). Donors were selected primarily on the basis of factors such as location, base flow index (BFI; Gustard *et al.* 1992) and regime similarity. For a few target stations in parts of the country where the network is sparse the choice of donors was restricted. Within the UK, Hydrometric Areas (HA) represent a group of connected catchments with one or multiple outlets to either the sea or a tidal estuary or alternatively a number of adjacent catchments of similar topography and separate tidal outlets (Marsh & Hannaford 2008). This study makes use of donors located both upstream and downstream of targets and in catchments belonging to the same or neighbouring HA. It is recognised that choice of donor catchment is likely to be an important influence, but the primary aim of this study is to determine how well various infilling methods perform given previously-defined donor catchments, rather than to consider the suitability of donors. Nevertheless, the potential impact that differing characteristics of donors can have on technique performance is considered in the interpretation of results.

Performance indices

Each method was tested by comparing the observed flow data from the target catchment against data simulated

Table 3 | Target station details, in ascending order of NRFA ID, and their corresponding primary and secondary donor stations

Target station NRFA ID	River (location)	Catchment area (km^2)	BFI	Mean flow ($m^3 s^{-1}$)	Period of record (used by study)	Primary donor NRFA ID	Secondary donor NRFA ID
7001	Findhorn (Shenachie)	415.6	0.36	13.96	1960–2008	7002	7004
15003	Tay (Caputh)	3210.0	0.64	140.27	1947–2008	15007	15006
21026	Tima Water (Deephope)	31.0	0.26	1.37	1973–2008	21017	21007
25003	Trout Beck (Moor House)	11.4	0.14	0.56	1957–2008	23009	76014
27071	Swale (Crakehill)	1363.0	0.46	20.73	1955–2008	27007	27034
28031	Manifold (Ilam)	148.5	0.54	3.52	1968–2008	28008	28046
29002	Great Eau (Claythorpe Mill)	77.4	0.89	0.68	1962–2007	29003	29001
33006	Wissey (Northwold)	274.5	0.82	1.83	1956–2007	33007	33019
33039	Bedford Ouse (Roxton)	1660.0	0.57	11.59	1972–2008	33037	33015
35003	Alde (Farnham)	63.9	0.37	0.31	1961–2008	35002	35013
38014	Salmon Brook (Edmonton)	20.5	0.29	0.16	1956–2008	38022	38021
38030	Beane (Hartham)	175.1	0.75	0.57	1979–2008	38004	33033
39101	Aldbourne (Ramsbury)	53.1	0.97	0.22	1982–2008	39077	39037
41023	Lavant (Graylingwell)	87.2	0.81	0.30	1970–2008	41015	42008
41029	Bull (Lealands)	40.8	0.38	0.45	1978–2008	41016	41003
43017	West Avon (Upavon)	84.6	0.71	0.69	1971–2008	53013	53002
46003	Dart (Austins Bridge)	247.6	0.52	11.20	1958–2008	46005	46008
54029	Teme (Knightsford Bridge)	1480.0	0.54	17.71	1970–2008	54008	55014
55029	Monnow (Grosmont)	354.0	0.5	5.97	1948–2008	56012	55013
63004	Ystwyth (Pont Llolwyn)	169.6	0.4	2.00	1984–2008	55008	63001
69017	Goyt (Marple Bridge)	183.0	0.53	3.75	1969–2008	69007	69015
74001	Duddon (Duddon Hall)	85.7	0.3	4.82	1968–2008	74007	74008
76003	Eamont (Udford)	396.2	0.52	15.51	1973–2008	76004	76015
85004	Luss Water (Luss)	35.3	0.28	2.63	1976–2008	86001	85003
93001	Carron (New Kelso)	137.8	0.26	10.85	1979–2008	4005	4006
96002	Naver (Apigill)	477.0	0.42	16.01	1977–2008	2002	3002

using the method, for the whole period-of-record rather than for any particular gap. Observed and simulated target flow series were compared using three commonly used indices, chosen according to the recommendations of studies which have explicitly compared such performance indicators and provided critical reviews of their utility (for example, Legates & McCabe 1999; Krause *et al.* 2005; Moriasi *et al.* 2007). These studies advocate the use of multiple performance indicators, due to the different strengths and limitations of the individual indices.

Nash-Sutcliffe model efficiency coefficient (NSE)

$$\mathrm{NSE} = 1 - \frac{\sum(\text{observed flow} - \text{estimated flow})^2}{\sum(\text{observed flow} - \text{mean}(\text{observed flows}))^2} \quad (1)$$

This statistic (NSE; Nash & Sutcliffe 1970) is extensively used within hydrology for evaluating model performance and, being standardised, is readily comparable across different catchments. The NSE provides an evaluation of the relative magnitude of the variance of the residuals compared to the variance of the observed flow data. Values can range from $-\infty$ to 1, with higher values implying greater accuracy and values below zero indicating that the simulated series is less accurate than if the mean of the observed series had been used. The NSE has been criticised for being overly influenced by higher flows and sensitive to errors in time-sequencing or when residuals are autocorrelated (e.g. Krause *et al*. 2005; McCuen *et al*. 2006; Beven 2012).

Root mean square error (RMSE)

$$\text{RMSE} = \sqrt{\left(\sum (\text{observed flow} - \text{estimated flow})^2\right)} \qquad (2)$$

RMSE is an absolute error measurement which is used to describe the difference between simulated and observed data in the unit of the variable, which aids in interpretation of the results. Both Legates & McCabe (1999) and Moriasi *et al.* (2007) recommend that measures of absolute error be used alongside dimensionless tests such as the NSE. Lower values indicate better performance, but comparing values between targets is limited since differing variance is not accounted for.

Percent bias (PBIAS)

$$\text{PBIAS} = \frac{\sum (\text{observed flow} - \text{estimated flow})}{\sum (\text{observed flows})} \times 100 \qquad (3)$$

The percent bias is useful in this study as, unlike the previous methods, it provides an indication of systematic bias in the simulated data. Positive (negative) values highlight consistent under(over)-estimation of target flows.

In addition to the above statistics, the means of the absolute residuals between observed and estimated flows were calculated for each target station. In order to judge the relative performance of infilling techniques against each other these means of residuals were compared using the non-parametric Wilcoxon test, to indicate whether a given technique generated estimated series with significantly lower means of residuals than those generated by other techniques. Using this measure, the percentage of cases where one technique significantly outperforms another can be compared.

RESULTS OF TECHNIQUE INTERCOMPARISON

Overall technique performance is illustrated by box and whisker plots of the NSE and PBIAS values derived from comparing the estimated and observed target series and grouped according to infilling technique (Figure 2), whilst bar charts of the NSE values for each target and technique contrast performance between the targets (Figure 3). The RMSE and NSE values yield very similar findings and hence the latter are presented in this paper since they represent standardised quantities.

The box and whisker plots of NSE values indicate that, for the vast majority of targets, all of the techniques generate estimated series with associated NSE values exceeding 0.5, albeit to varying degrees. Some techniques (MOVE.1 regression and catchment area scaling) feature outlying NSE values below zero, which suggests they are less applicable in some catchments. The strongest performing techniques are arguably the equipercentile and dual donor techniques, none of which have outliers below 0.5 and all of which have higher upper quartile, median and lower quartile values than the other techniques. Not only do these techniques therefore have broader applicability, but overall they produce estimated series which best replicate the observed target series. Catchment area scaling, on the other hand, emerges as a comparatively poorer technique for simulating daily time series, with the lowest upper quartile, median and lower quartile values and the greatest number of outliers.

The PBIAS values are generally of low magnitude, with the exception of those for the catchment area scaling technique, which is conspicuous for its tendency to consistently over- or under-estimate target flows. Techniques based upon log-transformed flows also exhibit bias to some extent. This can be connected to the failure of these techniques to maintain the mean of the observed target series in their estimates.

The bar charts (Figure 3) of NSE values expose some interesting disparities between technique performance for individual target stations. Whereas for some targets there is little distinction between the techniques (for example, 54029 and 85004), for others there is much greater

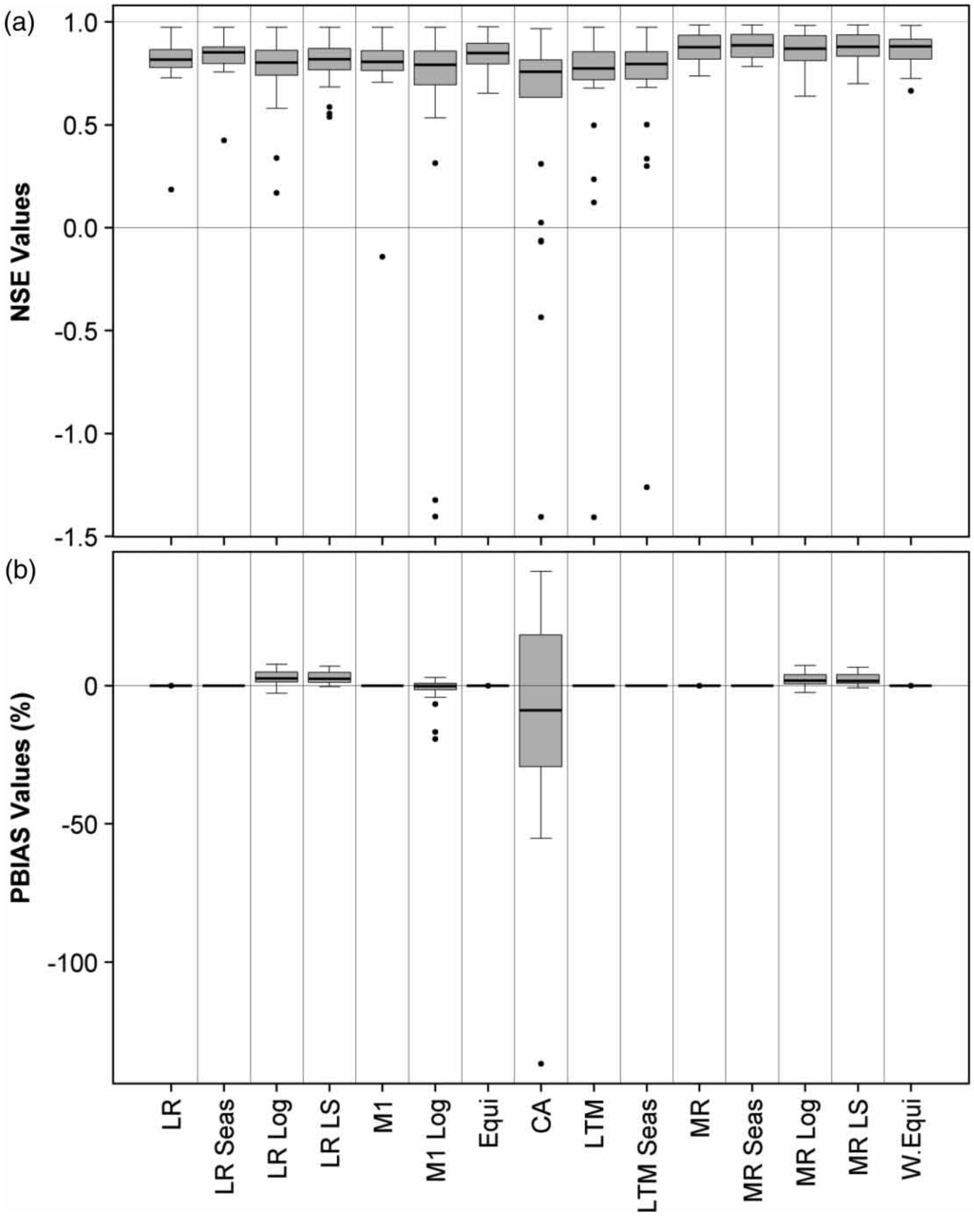

Figure 2 | Box and whisker plots of (a) NSE and (b) PBIAS values derived from comparing estimated and observed target series, grouped according to technique. Whiskers extend to the most extreme values which are no more than 1.5 multiplied by the interquartile range away from the box.

divergence (for example, 35003, 33039 and 76003). In certain cases, the dual donor techniques show distinctly higher simulation accuracy than the equivalent single donor techniques (for example, 33006 and 38014), endorsing the general argument of multiple donors being more capable of capturing the many influences affecting target flows. On many other occasions, however, the single and dual donor techniques yield similar performance, such that there is no marked advantage to including multiple donors.

The Wilcoxon significance testing results (Figure 4) further reinforce the above findings, with the equipercentile and dual donor techniques more frequently producing estimated target series with significantly lower means of

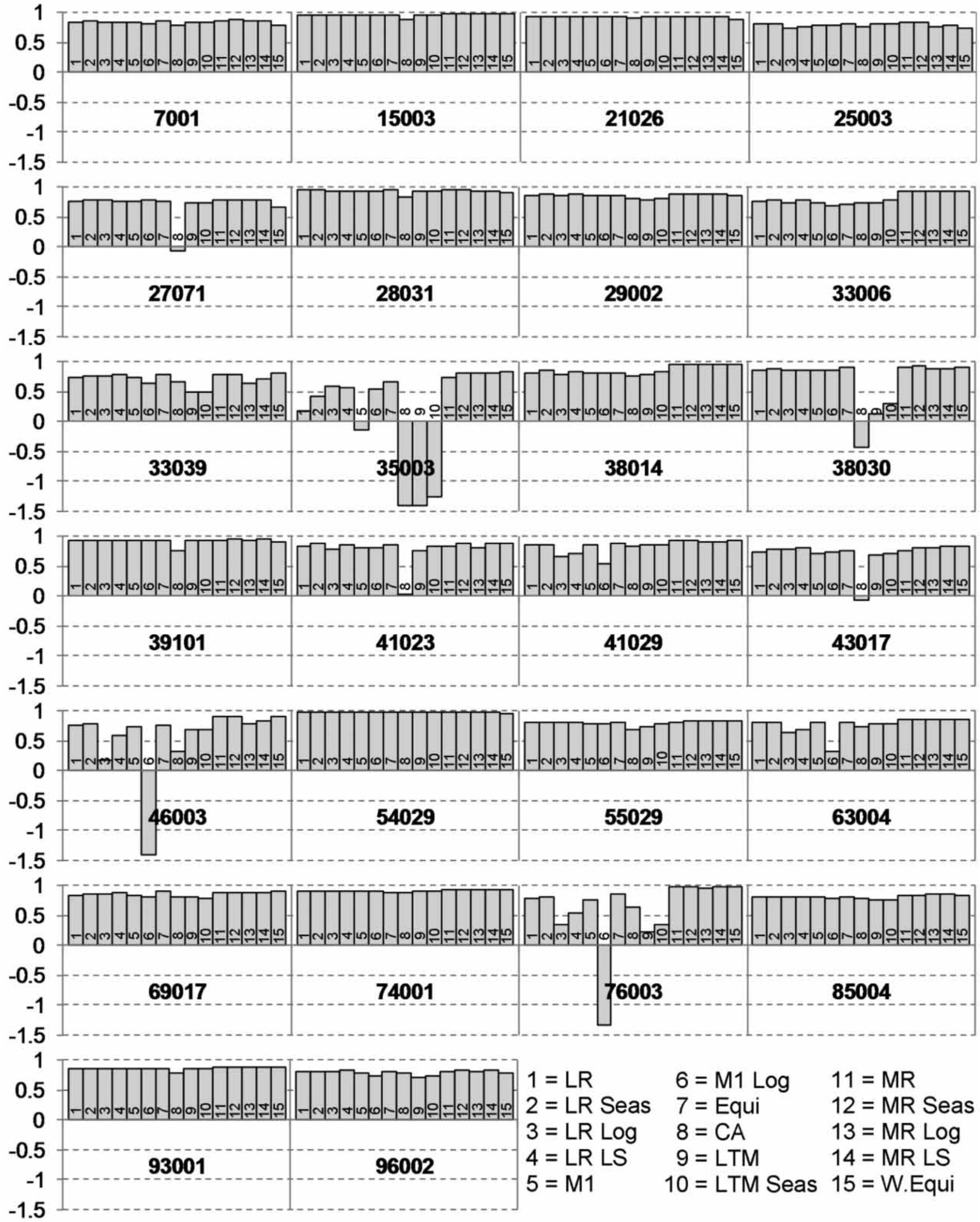

Figure 3 | Bar charts of NSE values derived from comparing estimated and observed target series.

residuals than the other techniques. All of the techniques outperform the catchment area scaling technique for the vast majority of the targets.

Comparing the ability of the chosen infilling techniques to estimate observed target flow series has revealed certain techniques which combine wide applicability with the

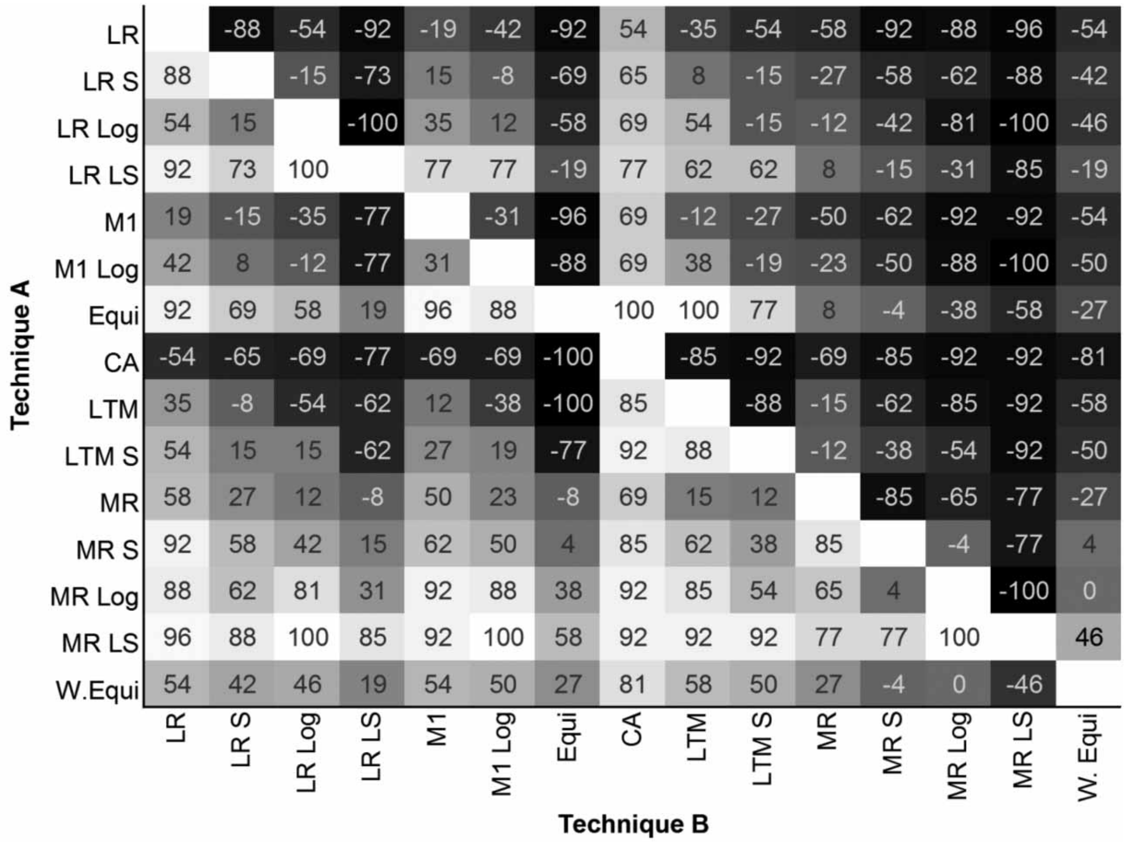

Figure 4 | Relative performance of different techniques as judged using Wilcoxon significance testing. Values at the intersection of technique A (*y*-axis) and technique B (*x*-axis) indicate the percentage more (positive values) or less (negative values) of stations for which technique A produced estimated flow series with significantly lower means of residuals (between the estimated and observed series) compared to those for estimates produced by technique B (at the 5% level). Colour-coded from black for 100% less cases to white for 100% more cases.

ability to outperform other techniques for specific target stations. This is a key outcome, highlighting the value of assessing a large sample of target stations. Despite its common usage it appears that, in the UK at least, catchment area scaling is essentially too simple to capture the influences affecting a target. Previous work has established that even closely related stations seldom exhibit a linear relationship with catchment area (Hughes & Smakhtin 1996). However, despite its poorer performance in daily flow infilling, catchment area scaling is widely used in hydrology (e.g. Shaw *et al.* 2010) and it has been shown to perform effectively in estimating summary hydrological characteristics (e.g. annual mean flow and annual peak flows). The limited range of hydroclimatic conditions in the UK mean that this finding cannot necessarily be generalised to other environments. The limited utility of scaling may reflect the spatial heterogeneity found in the UK but in regions with more homogeneous hydrological conditions the method may be more effective.

The results demonstrate that most techniques can perform competently across a broad spectrum of catchment types (see Table 1 for basic catchment characteristics). Indeed, the majority of techniques are shown to produce estimated series with NSE values exceeding 0.9 (Figure 3) for both large (15003) and small (21026) catchments, and across permeable, baseflow-dominated (39101) and impermeable, flashy (74001) regimes.

Results suggest that seasonally grouping flows prior to technique application enhances technique performance. NSE values are generally higher for seasonal based techniques (Figure 2(a)) and direct comparisons between the non-seasonal and seasonal applications of techniques show that in all cases the latter produces a significantly lower mean of residuals for over 85% more catchments than its non-seasonal equivalent (Figure 4). The same cannot be concluded for log-transforming flows, where non-transformed flow versions of techniques tend to produce estimated series with higher NSE values, but not significantly lower means of residuals. This most likely reflects the bias of the NSE statistic towards over-estimation of model performance at higher flows and under-estimation at lower flows, as identified by Krause *et al.* (2005), and its subsequent failure to capture the superior performance of the log-transformed versions of techniques when estimating lower magnitude flows (this issue is discussed in detail in one of the case studies presented below).

Calculation of the correlation coefficients between target and donor flows reveals donor station choice as a highly influential factor in technique performance. The five targets for which at least twelve techniques have associated NSE values exceeding 0.9 are those where observed flows have the highest correlations with the primary donor record (exceeding 0.95), whilst conversely the four targets for which at least twelve techniques have associated NSE values falling below 0.8 are those which have the lowest correlations with their primary donors (below 0.87). The links between superior technique performance and higher correlations between target and donor flows conform to general expectations since higher correlation coefficients indicate similar behaviour between flow regimes.

Technique performance also shows some correspondence to the relative locations of donors to their targets. For example, there are seven targets whose primary donors belong to different HA and in all of these cases none of the techniques generate estimated series with NSE values exceeding 0.9, which can be linked to lower correlations between flow series as a result of differing rainfall patterns and hydrological processes. With respect to dual donor techniques, there are eight targets whose donors share equivalent relative locations (for example, both are downstream on the same river as the target) and in only one of these cases the tested techniques yield an estimated series with an NSE value exceeding 0.9. On the contrary, there are four targets whose donors represent upstream and downstream versions of the same relative location (for example, both located on the same river as the target but one upstream and the other downstream), and for three of these the majority of techniques generate estimated series with NSE values exceeding 0.9.

For three of the targets (27071, 41023 and 43017, Figure 3), series estimated via catchment area scaling have distinctly lower NSE values than those estimated under all other techniques, in addition to the only PBIAS magnitudes exceeding 50%. The latter were negative in all three cases, signifying consistent over-estimation of target flows. This can be linked to the hydrogeology of the catchments involved, with the primary donor catchments generating proportionately higher runoff than those of the targets and thus leading to an over-estimation of target flows when employing this simple scaling technique.

PRACTICAL APPLICATIONS OF INFILLING METHODOLOGIES

The general conclusions that have been drawn thus far from the intercomparison between infilling techniques constitute a basis on which to develop broad infilling guidelines in future. However, there remain important issues which must be considered in applying such techniques in practice, which may limit the utility of the methods applied herein, and additional treatment of data may be required to address these issues prior to infilling. In this section, two important issues are discussed and illustrated using examples from the intercomparison sample of catchments (applying the full range of techniques used above) and recommendations are made for how these issues could be addressed in future. Subsequently two contrasting case studies are used as application examples, which demonstrate the best performing techniques, applied to new target catchments which have not been used in the intercomparison dataset.

Record inhomogeneity issues

The Salmon Brook at Edmonton (38014) gauges a small, impervious catchment in the south of the UK and originally comprised a compound broad-crested weir, known to be less effective than its 1980 flat V weir replacement (Marsh & Hannaford 2008). This hydrometric change manifests itself in a difference between pre-1980 and post-1979 data quality. Technique performance is shown to improve if the poorer quality data (pre-1980) is excluded before applying the infilling techniques (Table 4). This improvement is less

Table 4 | Comparison of NSE values derived from comparing estimated and observed series for targets 38014 and 33006, with datasets varied to exclude certain periods of record, in order to remove known inhomogeneities in the time series (for fuller explanation see text)

	NSE					
	Target 38014			Target 33006		
Infilling technique	Full datasets	Homogenised datasets	Change	Full datasets	Homogenised datasets	Change
LR	0.813	0.869	+0.055	0.745	0.834	+0.089
LR Seas	0.857	0.912	+0.055	0.782	0.859	+0.077
LR Log	0.777	0.852	+0.075	0.733	0.826	+0.093
LR LS	0.819	0.895	+0.076	0.777	0.852	+0.076
M1	0.804	0.864	+0.060	0.726	0.827	+0.101
M1 Log	0.796	0.861	+0.065	0.694	0.808	+0.115
Equi	0.809	0.865	+0.056	0.719	0.829	+0.110
CA	0.760	0.846	+0.085	0.730	0.821	+0.091
LTM	0.774	0.825	+0.050	0.739	0.816	+0.077
LTM Seas	0.822	0.873	+0.051	0.775	0.853	+0.078
MR	0.955	0.965	+0.010	0.933	0.933	+0.001
MR Seas	0.959	0.968	+0.009	0.935	0.937	+0.002
MR Log	0.948	0.957	+0.009	0.932	0.930	-0.002
MR LS	0.952	0.961	+0.009	0.935	0.934	-0.001
W. Equi	0.955	0.963	+0.009	0.924	0.929	+0.005

discernible for the dual donor techniques, a likely reflection of the fact that the primary donor record extends back to 1954 but the secondary donor record only started in 1971 (and thus a smaller number of poorer quality years are excluded when the approach is applied under the dual donor techniques). In addition, high NSE values are already associated with these techniques even before the poorer quality data are excluded (Figure 3).

An equivalent approach can be taken for station 33007 (the Nar at Marham), the primary donor associated with target 33006 (the Wissey at Northwold), a base flow dominated catchment in eastern England. Excluding pre-1987 data, to reflect the discontinuation of three groundwater abstractions in 1986 (Marsh & Hannaford 2008), leads to a noticeable improvement in the performance of the single donor techniques and, to a lesser degree, the majority of the dual donor techniques which, as before, reflect high NSE values even before excluding the early record data (Table 4, Figure 3).

The above examples demonstrate that removing a known inhomogeneity in a dataset, by carefully selecting the period of record considered, prior to applying infilling techniques can enhance technique performance. As well as the replacement or modification of gauging structures and changes concerning artificial influences, the homogeneity of UK flow records is affected by a host of other factors (e.g. instrumentation changes, catchment changes such as land-use) which are typically also major issues globally (Hannah *et al.* 2011). Assessing records for such changes and adapting them accordingly should therefore form an integral stage of infilling, which relies upon them being readily identifiable and underlines the necessity to maintain comprehensive metadata and user guidance alongside hydrometric records (Dixon 2010).

Estimating different magnitude flows

The Eamont at Udford station (76003) in north-west England gauges a catchment artificially influenced by controlled storage in lakes and reservoirs. Upstream donor stations were selected which have the same factors affecting runoff. For this target, the single donor techniques regressing log-transformed flows perform markedly more poorly overall than those regressing non-transformed flows (Figure 3). As would logically be expected, however, visual inspection of the estimated series suggests that log-transforming flows yields more reliable estimates of lower flows, despite less accuracy at higher flows (Figure 5). This suggests that combinations of techniques may offer the best solution to infilling flow data and a number of studies have previously advocated that a single technique is unlikely to be optimal for all occasions of missing data (for example, Gyau-Boakye & Schultz 1994; Hughes & Smakhtin 1996).

To further explore this finding and isolate techniques which consistently surpass others when estimating particular flow ranges, the accuracy of simulated target series was assessed according to primary donor flow magnitude. This reflects the practical application of infilling techniques, in that only donor flows will be available throughout a gap. Estimates were therefore grouped into three generalised classes of those relating to lower (Q95<Q≤Q65), medium (Q65<Q≤Q35) and higher (Q35<Q≤Q5) primary donor flows, thus ignoring the highest and lowest 5% of donor flow magnitudes since, as previously mentioned, estimating extreme flows is more challenging, and may not be appropriate due to the higher uncertainties associated with these data. In line with the findings outlined previously, where relevant, datasets were

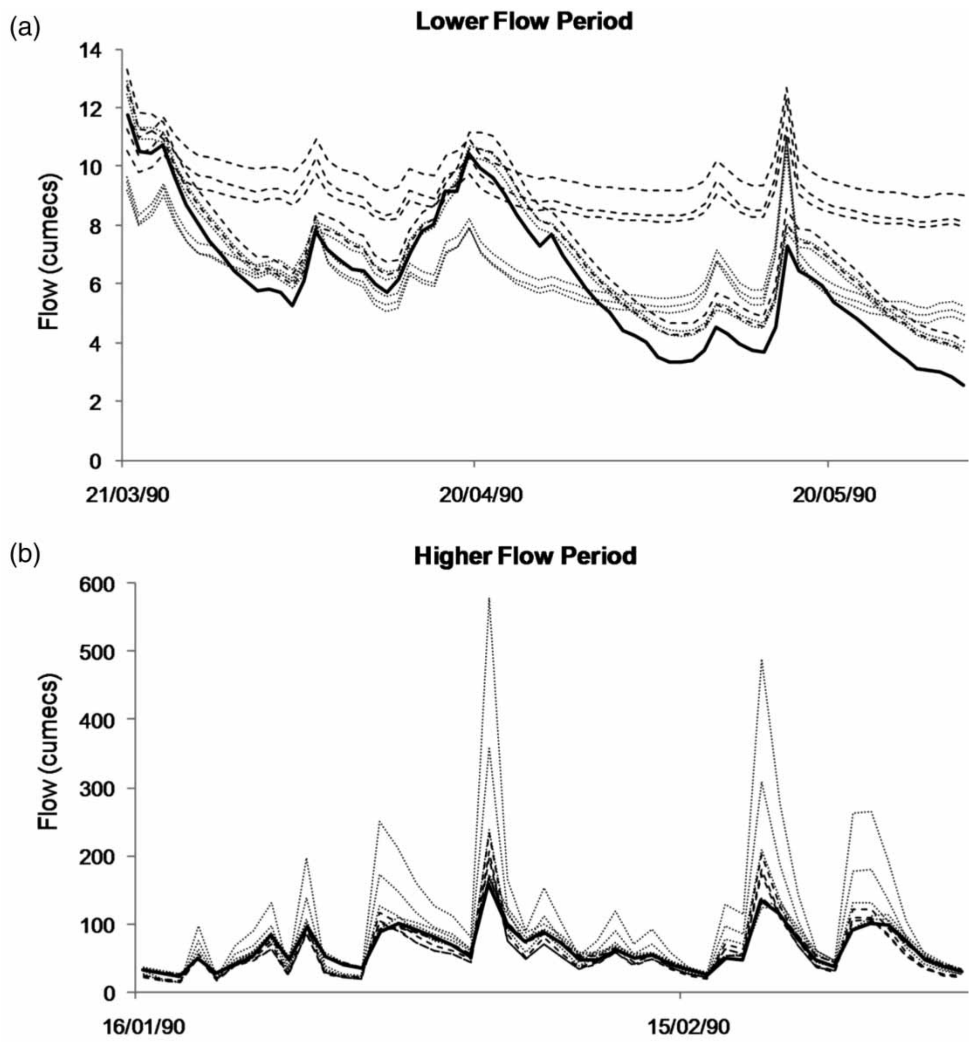

Figure 5 | Observed flows (solid black) for target 76003 during (a) a lower flow and (b) a higher flow period, with estimated series of dashed (dotted) lines corresponding to regression techniques under non-(log-)transformed flows.

first modified to develop more homogenous divisions of data.

Box and whisker plots of the NSE values associated to each group of estimates (Figure 6) show that, while variations are present, the relative general performance of the techniques reflects a similar pattern to those observed for the complete flow regime (Figure 2). The equipercentile and dual donor techniques maintain stronger performance across all magnitude groupings, whilst catchment area scaling is a poorer performing technique.

As expected for the lower flow magnitude class, the performance of the regression techniques based on log-transformed flows noticeably exceeds that of their counterparts based on non-transformed flows and, coupled with the equipercentile technique, these approaches demonstrate the strongest performance for this class (Figure 6(a)).

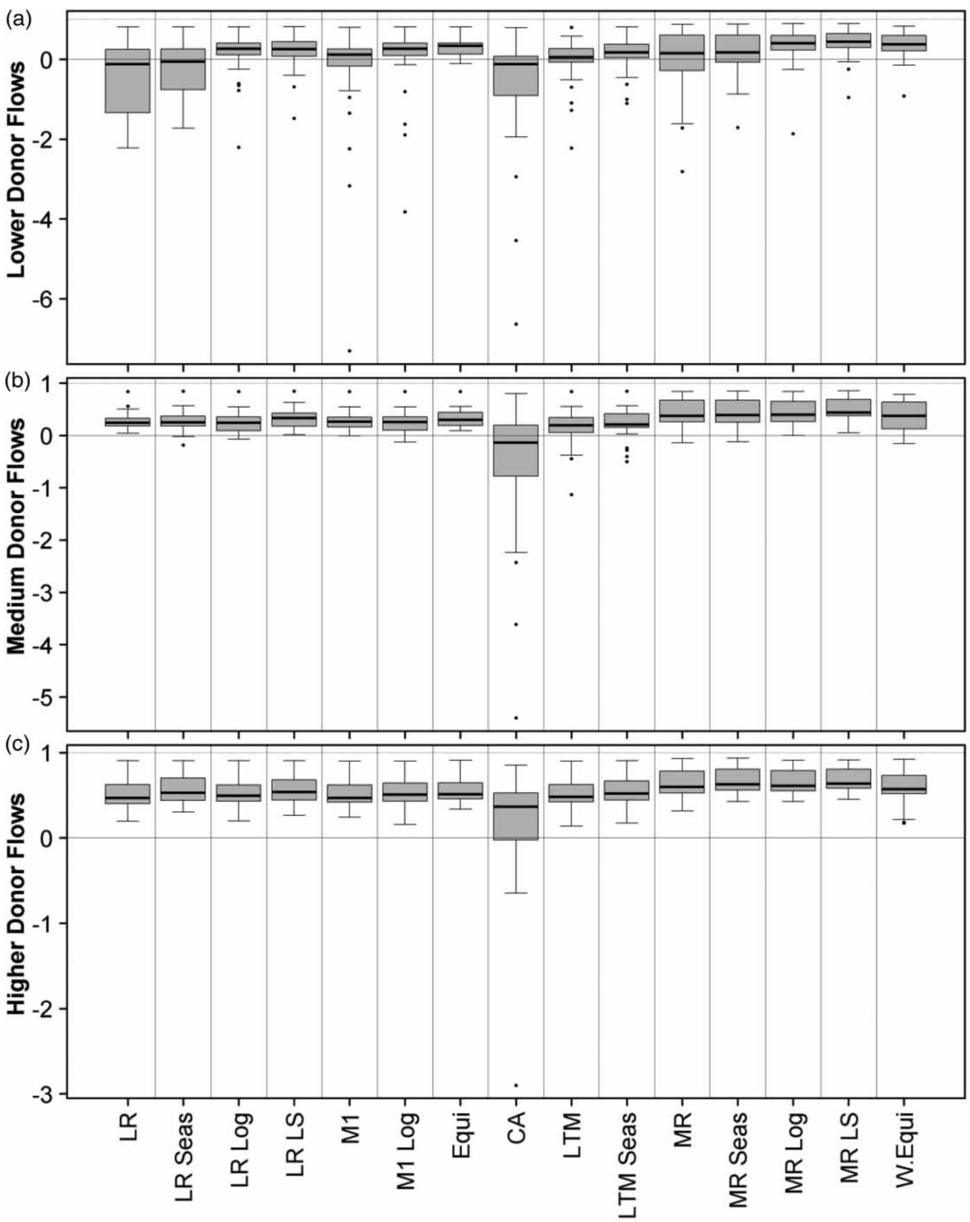

Figure 6 | Box and whisker plots of NSE values derived from comparing estimated and observed target series classified according to primary donor flow magnitudes, grouped according to technique. Whiskers extend to the most extreme values which are no more than 1.5 multiplied by the interquartile range away from the box.

It should be noted that the NSE values associated with the lower flow estimates for target 41023 have been excluded for clarity since they are the lowest for any technique and are extreme outliers in some cases. This is a consequence of flows at this target often falling to zero because of its ephemeral nature (due to permeable geology), behaviour which is not reflected in either donor and is therefore inherently difficult for statistical techniques to simulate.

The above result does not reverse for higher magnitude flows, with little difference evident between the performance of techniques based on non-transformed flows and their log-transformed counterparts, suggesting that the latter are the better overall choice.

Case study one: application of method in donor-rich environment

The Hore at Upper Hore Flume (54097) gauges a very wet, small and natural catchment, situated in the upper Severn basin in the Welsh uplands and belonging to the Plynlimon group of research catchments (Marc & Robinson 2007). Due to the density of instrumentation maintained in this area, other catchments of similar topography and flow regime can be selected as donors from within this group, namely the downstream Hore at Hore Flume (54092) as a primary donor and the neighbouring Tanllwyth catchment at Tanllwyth Flume (54090) as the secondary donor, both reflecting high correlations with the target of 0.994 and 0.990, respectively.

A 13-day long gap exists in the recent 2008 data of this station (Figure 7). It is reasonable to conclude that this period of record could be readily amenable to an infilling attempt, given the availability of good donors and the fact that, based on recorded rainfall patterns, nearby gauging stations and catchment response, the missing flows are likely to be mid-range. The best-performing single and dual donor techniques of the equipercentile approach (Equi) and multiple regression based on log-transformed

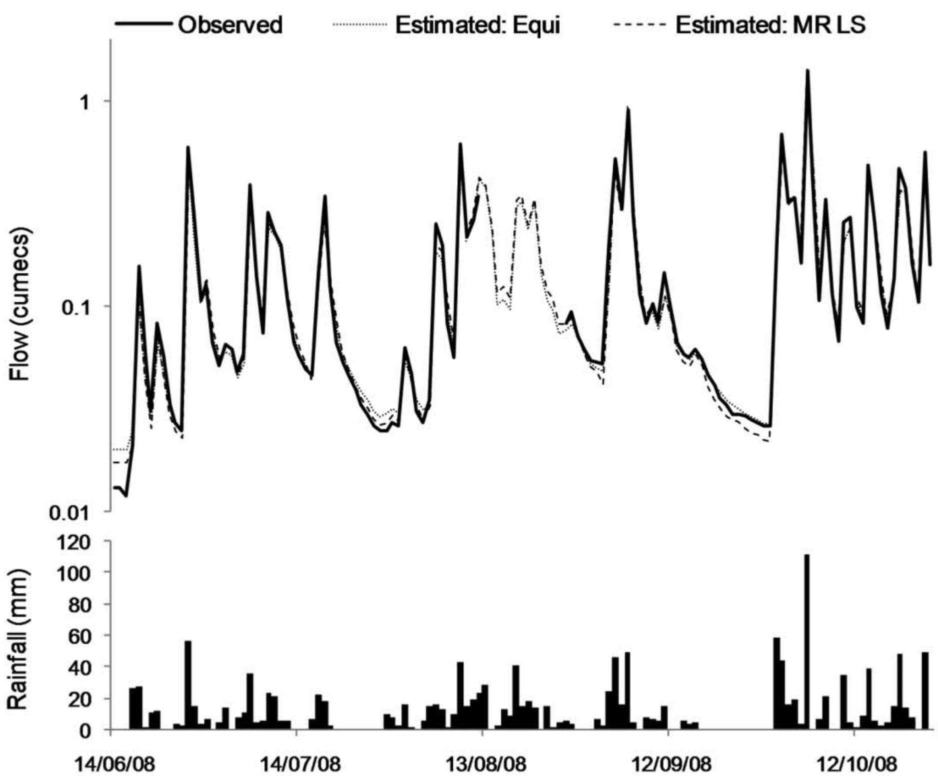

Figure 7 | Observed and estimated (under two different techniques) flow series for the Hore at Upper Hore Flume (54097), based upon donors of the Hore at Hore Flume (54092) and the Tanllwyth at Tanllwyth Flume (54090). Rainfall data from the Automatic Weather Station located within the Plynlimon group of research catchments at Carreg Wen is shown for verification purposes.

and seasonally grouped flows (MR LS), as judged by the results of this study, were therefore applied to generate estimates of flow to infill the gap (Figure 7). Both techniques yield estimated series which are similar to the available observed flows around the time of the gap. While this paper deliberately excludes infilling techniques that are dependent on data other than river flows, other hydrometeorological observations such as rainfall records may provide a useful verification of results. In the case of this small responsive catchment, rainfall data from a closely located gauge suggest that the estimated flow pattern reflects rainfall recorded throughout the gap. This example therefore illustrates how the simple infilling techniques presented by this study, coupled with local donor station data, can successfully generate infill estimates that can be adopted with some degree of confidence in order to improve the completeness and utility of a flow record.

Case study two: application in a donor-poor environment

The Crimple at Burn Bridge (27051) is a small catchment in northern England that forms part of the UK Benchmark Network of natural catchments, often used within climate change detection studies (for example, Hannaford & Marsh 2008). Consequently, it is of particular importance that its record be as complete as possible, to allow the calculation of long-term trends and summary statistics. While the record is complete post-1982, infilling a gap during the earlier part of the record in 1975 poses a challenge due to the difficulty in identifying suitable donor stations, with the records of many nearby stations either commencing after 1975 or being subject to heavy reservoir influences (Figure 8). Problems such as these are widespread in the UK, where the degree of artificial influence on many flow regimes is such that practitioners will often be faced with the question of whether catchments with varying factors affecting runoff can usefully provide information transferable to neighbouring catchments under certain flow conditions.

Data-infilling techniques were applied using various combinations of the potential donors and coupled with knowledge of any differing influences on the flow regimes

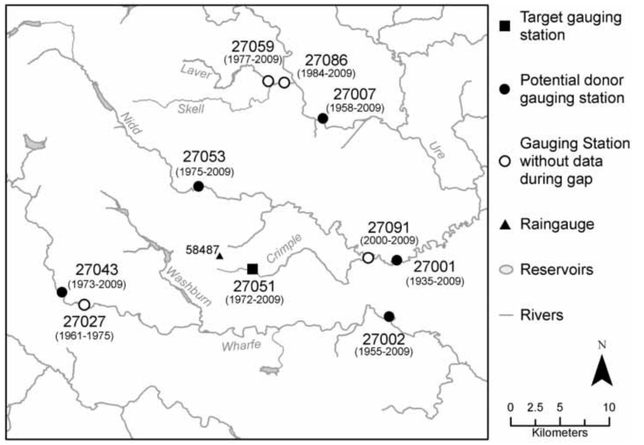

Figure 8 | Locations of target station 27051 (the Crimple at Burn Bridge) and nearby potential donors. The period of river flow record held on the NRFA is shown in brackets for each gauging station.

to assess suitability. Estimates derived via the equipercentile technique using a donor from a neighbouring catchment, the Nidd at Birstwith (27053; Figure 8) are shown in Figure 9. While the upper catchment is influenced by reservoir storage control, the Nidd is subject to the same meteorological controls at the target station and the degree of regulation is such that much of the natural flow pattern is maintained. Consequently, records from the station may have donor utility under certain flow conditions. Results show that the estimated time series captures the general variability of the target flows and, as expected, during some winter periods of the record provide an analogue of observed flows. However, around the time of the target record gap there are many notable discrepancies, timing shifts and consistent over-estimation in the estimated series due to the differing controls on flow. In this case, the infilled data would not be considered representative of flow behaviour during the missing sequence.

This example thus illustrates the importance of donor selection and knowledge of the factors affecting flow regimes and shows that even a technique which generally performs well cannot be expected to provide reliable infill estimates in cases where appropriate donors are limited. Such issues highlight some of the dangers associated with any fully automated application of the data-infilling techniques and show that manual appraisal of derived estimates are an imperative step in determining whether infilled data can be adopted.

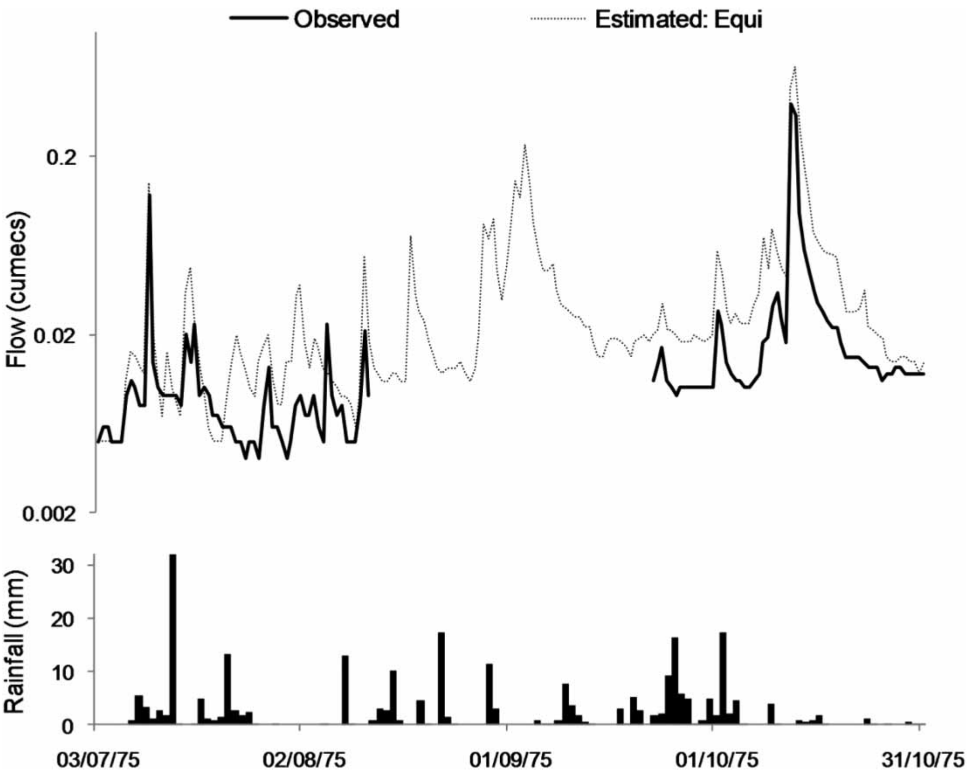

Figure 9 | Observed and estimated (under the equipercentile technique) flow series for the Crimple at Burn Bridge (27051), based upon a donor of the Nidd at Birstwith (27053). Rainfall data from the Met Office raingauge at Ten Acres Reservoir (National Raingauge 58487).

OTHER CONSIDERATIONS FOR DATA INFILLING

This study has provided an evaluation of existing techniques in terms of their performance in estimating observed flows and their versatility in application, and has also outlined some of the practical issues in applying the techniques. There are undoubtedly additional practical issues to consider and future work will focus on application of the techniques under specific hydrometric situations to assist in the development of operational guidance for practitioners. For example, the current study applies methods to simulate whole time series but, in practice, depending on the cause of the missing data, the period either side of a gap in the target series may also guide estimation. Use of infilling techniques may also result in discontinuities between the infilled data and observed data either side of the gap; any future guidance must consider how this could be addressed. By focusing on whole time series, this study has not addressed the practical question of gap duration or the location of gaps in the flow regime (although consideration was given to applying methods to separate magnitude classes within the flow regime). Future work will seek to address guidelines for gap-duration or location, e.g. specifying maximum gap length appropriate for infilling. In this context, other time series characteristics which were not assessed in this study – such as the capacity of infilling methods to reproduce autocorrelation structure and long-term persistence (e.g. Koutsoyiannis 2002) – may be especially important as gap length increases, and should be considered in future studies. Time lag between target and donor flows was not addressed as it was not important for this daily dataset, due to the relative size and rapid runoff of most UK catchments, but may be an important component of any future guidance.

Within the UK, the findings of this study will support the development of general infilling guidelines appropriate to a wide range of flow regimes, while also presenting practitioners with a selection of targeted infilling techniques, with local hydrological conditions and the hydrometric experience of the measuring agencies guiding the ultimate choice of method and its application. There are undoubtedly many instances where an infill would not be appropriate, especially when suitable donor stations cannot be found due to network sparseness, heavy artificial influences or hydrometric inadequacies. Even if a good donor is available, other factors could limit the applicability of these methods. Infilling during flood periods is likely to be subject to huge uncertainties but, arguably, these may be the circumstances when recovering missing data would be of greatest practical importance. From the standpoint of flood frequency analysis, some form of estimate would be preferable to having no knowledge of event magnitude. However the methods used in this study are unlikely to be as useful as traditional methods for estimating peak flow using reconstructed levels and hydraulic theory (e.g. Herschy 2009), hydraulic models or rainfall-runoff methods.

Finally, whilst statistical data transfer techniques are an important tool to aid the infilling of missing or erroneous observational records, it is important to recognise that the resulting infilled data only provide an estimate of river flow during the period in question, and should be identified using metadata flags and comments to guide users.

CONCLUSIONS

Complete river flow records are vitally important to water resources management but obtaining such series can be very difficult, given the many means by which gaps can arise in observed data. Simple infilling techniques, with the potential to derive reliable estimates for a broad range of flow regimes, could therefore find wide utility in an operational setting where more complex catchment modelling is not practical. Systematic appraisals of techniques, such as the one presented in this paper, are an important step in promoting a consistent methodology for minimising record gaps.

This study is distinctive in its assessment of multiple techniques – 15 in total, all relying upon single or dual donor station data transfer – according to their ability to generate estimated flow series for a broad sample of 26 representative UK gauging stations. Findings demonstrate the alliance of superior technique performance with a strong correlation between target and donor flows (linked to relative donor station location) and the improvement in performance associated with applying techniques to more homogeneous datasets. The aim of the study has not been to isolate a single optimal technique, but to instead explore the range of applicability and general performance of each of the techniques. Key findings suggest that, overall, the equipercentile and dual donor techniques have demonstrated their potential to derive more accurate infill estimates, whilst catchment area scaling has been conspicuous through its poorer performance.

Outside the sphere of operational hydrometry, adopting a uniform, repeatable approach towards infilling gaps in river flow data promises many possible advantages to scientists and practitioners both within the UK and internationally. Nevertheless, it must be emphasised that, despite the aptitude of simple infilling techniques to generate reasonable flow estimates, as illustrated by the examples presented within this study, maximising the quality and completeness of observed river flow data should be the first and foremost priority.

ACKNOWLEDGEMENTS

The authors thank their colleague Terry Marsh at the Centre for Ecology & Hydrology for his assistance with this study, and Mark Robinson for reviewing the paper. The authors also wish to thank the two anonymous reviewers for their assistance in improving the manuscript. All data were sourced from the National River Flow Archive at http://www.ceh.ac.uk/data/nrfa.

REFERENCES

Amisigo, B. A. & van de Giesen, N. C. 2005 Using a spatio-temporal dynamic state-space model with the EM algorithm to patch gaps in daily riverflow series. *Hydrol. Earth Syst. Sci.* **9**, 209–224.

Beven, K. J. 2000 Uniqueness of place and process representations in hydrological modelling. *Hydrol. Earth Syst. Sci.* **4**, 203–213.

Beven, K. J. 2012 *Rainfall-Runoff Modelling: The Primer*. 2nd edition, Wiley, Chichester, 457 pp.

Dixon, H. 2010 Managing national hydrometric data: from data to information. In: *Global Change – Facing Risks and Threats to Water Resources* (E. Servat, S. Demuth, A. Dezetter & T. Daniell, eds). Proceedings of the Sixth World FRIEND

Conference, Fez, Morocco. October 2010. IAHS Publication, Wallingford, UK, 340, pp. 451–458.

Elshorbagy, A. A., Panu, U. S. & Simonovic, S. P. 2000 Group-based estimation of missing hydrological data: II. Application to streamflows. *Hydrol. Sci. J.* **45**, 867–880.

Gustard, A., Bullock, A. & Dixon, J. M. 1992 Low flow estimation in the United Kingdom. Institute of Hydrology Report No. 108.

Gyau-Boakye, P. & Schultz, G. A. 1994 Filling gaps in runoff time-series in West-Africa. *Hydrol. Sci. J.* **39**, 621–636.

Hannaford, J. & Marsh, T. J. 2008 High-flow and flood trends in a network of undisturbed catchments in the UK. *Int. J. Climatol.* **28**, 1325–1338.

Hannah, D., Demuth, S., van Lanen, H., Looser, U., Prudhomme, C., Rees, H., Stahl, K. & Tallaksen, L. 2011 Large-scale river flow archives: importance, current status and future needs. *Hydrol. Process*. **25**, 1191–1200.

Herschy, R. W. 2009 *Streamflow Measurement*. 3rd edition, Taylor and Francis, Abingdon, 507 pp.

Hirsch, R. M. 1979 An evaluation on some record reconstruction techniques. *Water Resour. Res.* **18**, 1081–1088.

Hirsch, R. M. 1982 A comparison of four streamflow record extension techniques. *Water Resour. Res.* **15**, 1781–1790.

Hughes, D. A. & Smakhtin, V. 1996 Daily flow time series patching or extension: a spatial interpolation approach based on flow duration curves. *Hydrol. Sci. J.* **41**, 851–871.

Ilunga, M. & Stephenson, D. 2005 Infilling streamflow data using feed-forward back-propagation (BP) artificial neural networks. Application of standard BP and pseudo Mac Laurin power series BP techniques. *Water SA* **31**, 171–176.

Khalil, M., Panu, U. S. & Lennox, W. C. 2001 Groups and neural networks based streamflow data infilling procedures. *J. Hydrol.* **241**, 153–176.

Kottegoda, N. T. & Elgy, J. 1977 Infilling missing flow data. In: *Modelling Hydrologic Processes* (H. J. Morel-Seytoux, ed.), Water Resources Publications Littleton, CO, pp. 60–73.

Koutsoyiannis, D. 2002 The Hurst phenomenon and fractional guassian noise made easy. *Hydrol. Sci. J.* **47**, 573–596.

Krause, P., Boyle, D. P. & Bäse, F. 2005 Comparison of different efficiency criteria for hydrological model assessment. *Adv. Geosci.* **5**, 89–97.

Legates, D. R. & McCabe Jr, G. J. 1999 Evaluating the use of 'goodness-of-fit' measures in hydrologic and hydroclimatic model validation. *Water Resour. Res.* **35**, 233–241.

Marc, V. & Robinson, M. 2007 The long-term water balance (1972–2004) of upland forestry and grassland at Plynlimon, mid-Wales. *Hydrol. Earth Syst. Sci.* **11**, 40–60.

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

Marsh, T. J. & Hannaford, J. (eds) 2008 *UK Hydrometric Register. Hydrological data UK series*. Centre for Ecology & Hydrology, Wallingford, UK.

McIntyre, N., Lee, H. & Wheater, H. 2005 Ensemble predictions of runoff in gauged catchments. *Water Resour. Res.* **41**, pW12434.

Merz, R. & Blöschl, G. 2005 Flood frequency regionalisation – spatial proximity vs. catchment attributes. *J. Hydrol.* **302**, 283–306.

McCuen, R. H., Knight, Z. & Cutter, A. G. 2006 Evaluation of the Nash-Sutcliffe efficiency index. *J. Hydrol. Eng.* **11**, 597–602.

Moriasi, D. N., Arnold, J. G., Van Liew, M. W., Bingner, R. L., Harmel, R. D. & Veith, T. L. 2007 Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. *T. ASABE* **50**, 885–900.

Nash, J. E. & Sutcliffe, J. V. 1970 River flow forecasting through conceptual models. Part 1: A discussion of principles. *J. Hydrol.* **10**, 282–290.

Raman, H., Mohan, S. & Padalinathan, P. 1995 Models for extending streamflow data: a case study. *Hydrol. Sci. J.* **40**, 381–393.

Rees, G. 2008 Hydrological data. In: *Manual on Low-flow Estimation and Prediction. Operational Hydrology Report No. 50.* (A. Gustard & S. Demuth, eds). World Meteorological Organisation, Geneva, Switzerland.

Smakhtin, V. Y. & Masse, B. 2000 Continuous daily hydrograph simulation using duration curves of a precipitation index. *Hydrol. Process.* **14**, 1083–1100.

Shaw, E. M., Beven, K. J., Chappell, N. A. & Lamb, R. 2010 *Hydrology in Practice*. Fourth edition, Spon Press, Abingdon, UK, 543 pp.

Wagener, T., Sivapalan, M., Troch, P. & Woods, R. 2007 Catchment classification and hydrologic similarity. *Geogr. Compass* **1**, 901–931.

Wallis, J. R., Lettenmaier, D. P. & Wood, E. F. 1991 A daily hydroclimatological data set for the Continental United States. *Water Resour. Res.* **27**, 1657–1663.

World Meteorological Organization (WMO). 2008 *Guide to Hydrological Practices*. Sixth edition, WMO-No. 168, Geneva, Switzerland.

Yadav, M., Wagener, T. & Gupta, H. 2007 Regionalization of constraints on expected watershed response for improved prediction in ungauged basins. *Adv. Water. Resour.* **30**, 1756–1774.

First received 16 November 2010; accepted in revised form 10 October 2011. Available online 12 April 2012