Flow classification provides a statistically robust method of defining an expected range of variability for flow metrics describing frequency, magnitude, duration and timing of events. Here, we characterised reference mean daily flows for 1950–1999 for all 5838 quinary catchments of South Africa based on 150 metrics. Using a two-tiered approach, where sub-catchments were classified into similar flow types using principal components and cluster analyses, we defined 6 to 12 flow types for each of 8 hydrological regions reflecting rainfall seasonality. Redundancy between variables was 87% on average, so that site variability could be accounted for using 8–28 metrics. In general, flow volume metrics accounted for Axis 1 variability, while coefficients of dispersion had 1.8 times less leverage in Axis 2. With the incorporation into a spatial product and an associated database, this study provides a basis for defining statistically robust reference flow conditions for multiple flow metrics, against which current observed flows at specific sites may be compared.

INTRODUCTION

Flow classification is a statistical process of stratifying natural variation in measured characteristics among a population of streams and rivers to delineate river types that are similar in terms of hydrologic...features (Poff et al., 2010 p. 153). The degree of deviation is best represented by measuring statistical departure from reference condition sensu Stoddard et al. (2006). Since flow regimes vary geographically in response to climate and catchment characteristics (geology, stream order, topography, land cover) (Poff and Zimmerman, 2010), flow classifications are appropriate at a regional scale, and are useful because they integrate landscape features into a single measurable index (Gordon et al., 1994). Unlike agglomerative approaches that make use of techniques such as duration curves, reductionist classification approaches make use of indices (= metrics) that focus on state and threshold values using descriptive statistics, and attempt to understand the links between timing, duration and magnitudes of different system states. Numerous metrics of flow types can be readily calculated (for example: coefficient of intra-annual variation; number of zero-flow days per year; and Colwell's (1974) indices of predictability and constancy), which provide an indication of predictability and perenniality, while average monthly flows provide a measure of seasonality (Haines et al., 1988; Poff and Ward, 1989). Olden and Poff (2003) assessed and grouped 171 hydrologic indices into 5 categories representing biologically relevant streamflow attributes, based on the work of Richter et al. (1996) who developed the 'Indicators of Hydrological Alteration', an approach which analyses flow time-series for ecologically relevant flow measures, based on 32 parameters which are divided into 5 groups relating to magnitude, timing, duration and frequency of ecologically significant events. Examples of the use of flow metrics to classify flow types include streamflow indices to classify Victorian (Australia) streams, where 16 variables were used to classify streams into 5 groups (Hughes and James, 1989); 78 streams in the United States were classified into 9 types based on 15 variables (based on, inter alia, floods, predictability, zero flows and seasonality) (Poff and Ward, 1989); and the use of 120 metrics and a classification undertaken for Australian rivers (Kennard et al., 2010).

Increasingly over the past few decades, the detrimental impacts of changes to flow regimes have been recognised. Causes of this are multifarious and include, inter alia, flow reductions due to surface water abstractions or changes in basflow due to abstractions of groundwater, changes in runoff due to inappropriate catchment land use (increases or decreases), fundamental changes to downstream river hydrographs as a result of upstream impoundments (Palmer and O’Keeffe, 1989), and impacts to donor systems (reductions) and receiving system (increases) due to inter-basin transfer schemes (Snaddon and Davies, 1998).

Changes to flows are typically reflected in biological responses (Jackson et al., 2007). Flow patterns fundamentally impact the types and distributions of aquatic species within river systems (De Moor, 2002; Bunn and Arthington, 2010; Poff and Zimmerman, 2010), with streamflow perceived as a 'master variable' shaping many fundamental ecological characteristics of riverine ecosystems (Poff and Zimmerman, 2010). This occurs both through direct impacts on species based on their hydraulic preferences and tolerances, as well as indirectly because of the impacts of flows on geomorphology that create the abiotic habitat templates for species (Statzner and Higler, 1986; Poff et al., 2010). For example, in South Africa, this has been studied over the past 30 years in relation to blackfly problems on the Great Fish River (O’Keeffe and De Moor, 1988; Rivers-Moore et al., 2007, 2008), and may even translate into considerable economic losses, as evidenced by similar issues, albeit on a greater scale, along the middle and lower reaches of the Orange River (Rivers-Moore et al., 2014). Not only is too
little flow deleterious to river health, but too much flow is also problematic, and the Great Fish River in South Africa's Eastern Cape Province is cited as a classical example of a permanently altered system as a consequence of this (O’Keefe and De Moor, 1988). Poff and Zimmerman (2010) highlight that 92% of a total of 165 papers which they reviewed reported negative impacts on ecological metrics based on flow alteration. An explicit spatial product resulting from regional flow type classifications could enable researchers developing generalisations about hydrology–ecology interactions on the basis that responses to natural flow regime characteristics are likely to vary between flow types (Kennard et al., 2010).

Flow classification serves two important purposes in the environmental flows determination process: firstly, relationships between ecological metrics and flow alteration can be developed for an entire river type based on data from a limited set of rivers within each region; and secondly, it facilitates more efficient biological monitoring and research design (Poff et al., 2010). In South Africa, broad-scale aquatic biodiversity patterns are already well recognised: Harrison’s (1959, cited in De Moor, 2006) 12 hydro-biological regions; distinct upland versus lowland aquatic macro-invertebrate assemblages as shown by Dallas (2004); and aquatic zones (such as Level 1 Ecoregions; Kleyhans et al., 2005). Streamflow classifications are a logical extension of such regional perspectives, and are useful because they integrate landscape features into a single measurable index (Gordon et al., 1994). Underpinning any regional environmental flow assessment is a hydrological classification based explicitly on flow data (Kennard et al., 2010; Olden et al., 2012).

The need to classify rivers in South Africa first arose in the 1960s, and re-emerged as an issue in the mid-1980s (Eekhout, 1997). Joubert and Hurly (1994) classified South African rivers into three broad flow categories based on seasonality, variability and perenniality. Ten flow classes incorporating this logic were subsequently included into a multi-criteria analysis framework for defining environmental flows (Brown and Joubert, 2003) within an environmental flow holistic framework that has been extensively used in multiple countries (Tharme, 2003). Hughes and Hannart (2003) classified flows into classes according to a hydrological index (HI) developed for South African rivers and applied at a quaternary catchment scale. The HI is based on the ratio of monthly coefficients of variation for flows to a baseflow index, to provide a unitless value that reflects seasonality and perenniality of rivers at a quaternary catchment scale. Monthly coefficients of variability typically reflect rainfall patterns, and give some indication of seasonality, while the baseflow coefficient is the proportion of total flow that occurs as baseflow, and provides a measure of short-term variability by reflecting how peak flows are concentrated or spread throughout the year. This approach has been used as a desktop approach for regional environmental flow assessment. In spite of such a body of applied research and progress in defining environmental flows, gaps nevertheless exist in terms of a national database of accessible reference flow metrics that also have ecological applications. In this study, our aim was to develop a regional flow type classification using regionally specific metrics, as a spatial tool to assist in defining reference flows.

METHODS

Our flow type classification process followed equivalent steps to those described by Poff et al. (2010); viz. obtain flow data; describe the flow time-series in terms of metrics; undertake regional redundancy analyses; classify flow types per quinary sub-catchment, and assign to a spatial framework. For South African flow conditions, Taylor et al. (2003) recommends using time-series of either observed or estimated daily natural flows with at least 20 years of reasonably complete time-series of mean daily flow rates for reliable conclusions to be drawn. For the flow data, we used simulated data based on Acock’s veld types for baseline current conditions (1950–1999) from 5838 quinary catchments. These data were selected as the basis for defining ‘reference’ flow types, and form part of a database developed in previous studies (Schulze, 2010; Mahery et al., 2013), based on widely verified simulated flows using the process-based agro-hydrological daily time step ACRU model developed for South African hydrological conditions (Schulze, 1995; Smithers and Schulze, 2004). The first 2 years of model output were deleted in order to obtain realistic baseflow and soil water stores, as per the recommendations of Taylor (2006), so that time-series data spanned the period 1 October 1952–31 December 1999.

Next, flow time-series were described in terms of metrics. Each flow data file was analysed using non-parametric statistics, using the IHA approach (Indicators of Hydrologic Alteration; Richter et al., 1996; Mathews and Richter, 2007), and following a classification undertaken for Australian rivers using 120 metrics (Kennard et al., 2010) that is relevant because of the known similarities in hydrological regimes between Australia and South Africa (Chiew et al., 1995). To test the validity of using the simulated flows to define reference flow types, we compared simulated flows and their associated metrics with reliable concurrent flow data. We selected 4 gauging weirs previously confirmed as having reliable, long-term data, i.e. > 35 years of data tested for stationarity, consistency, homogeneity and absence of trend and characterised as ‘reasonably natural’ streamflow by Taylor (2006). The verification process involved visual assessments of concurrent observed and simulated mean daily flow data using time-series plots; correlations using linear regression analyses of observed versus simulated mean daily flows; double mass plots of cumulative observed versus simulated mean daily flows; and correlations between IHA metrics for observed versus daily flows, using simple linear regression.

To account for regional differences in flow metric redundancies that were highlighted by Taylor (2006), and which would be due to spatio-temporal variation in orographic rainfall patterns, we adopted a two-tiered approach to regional flow type classification. For the first level, we defined hydrological regions using the rainfall regions of Schulze and Maharaj (2007) to account for ecological similarity at a biogeographical scale, and water management areas (WMAs) for South Africa for ease of applicability.

For the next step, flow metrics were iteratively screened for multi-collinearity for each region to reduce data redundancy (Olden and Poff, 2003). The maximal variable set per region was refined using a combination of principal component (PC) analysis (PCA) (McCune and Meford, 2011) and spreadsheet correlations to eliminate variables with either no or poor correlations with PC Axes 1 and 2, or where variables with an R² correlation of ≤50% (Pearson test) for PC Axes 1 and 2 were eliminated. A further PCA iteration was run, with variables used in this step compared in a spreadsheet using a correlation matrix. Variables that showed a high degree of collinearity were identified, and the variable with the highest eigenvalue from the PCA was selected for inclusion in the optimal variable matrix. Variables with high correlations but lower eigenvalues were deleted, as these did not add to the explanatory power of the PCA. The final PCA was run using the optimal matrix of variables. The cumulative variance accounted for in the first three PC axes was noted for successive iterations, with the expectation that the cumulative variation accounted for in the PCA based on the maximal dataset would be less than successive iterations. However, too much pruning of variables was also to be avoided, and this was indicated by a decline in the explained variance.
Site groups and river types for each region were defined using the optimal dataset resulting from the redundancy analysis. Principal component (correlation matrix method) and cluster (Euclidean distance measure; group averaging technique) analyses (McCune and Mefford, 2011) were used to define flow type groups per hydrological region. The quinary vector coverage was attributed by flow type classes, with an associated database file for annual, median and coefficient of dispersion metrics linking to each quinary catchment.

By way of examples in the use of the spatial database of metrics and flow types, we applied the metrics for different flow types in three different ways. Firstly, expected reference seasonal variation was calculated for the Thukela catchment, with monthly median flows shown using box-and-whisker plots. Secondly, the range of predictability in flows across all flow types within Region 1 (Western Cape winter rainfall) was calculated using box-and-whisker plots. Finally, we used mean monthly flow data from a gauging weir (H1H009) downstream of the Stettynskloof Dam on the Holsloot River, a tributary of the Breede River in the Western Cape. Mean monthly flows for the hydrological year 2014/15 were plotted against a reference curve for the corresponding flow type that included a 95% confidence envelope.

**RESULTS**

Validity of using simulated flows to define reference flow types

A total of 150 metrics were derived for each quinary catchment: 7 metrics for annual flows; 69 metrics for median values describing the full time-series including monthly statistics, and 74 associated metrics for coefficients of dispersion. Whereas comparisons of observed and simulated flow data showed agreement in terms of pulses when compared as visual plots, correlations between observed and simulated mean daily flows ranged from poor to fairly strong ($R^2$ of 0.05–0.72; Table 1). It is most likely that these differences are primarily a consequence of changes in timing and magnitude of flow events in response to degrees of departure from ‘natural’ land cover, which did not always show an ideal homogenous relationship. Such inconsistencies became less apparent in the comparison of IHA metrics based on more than 40 years of observed versus simulated flows, with all $R^2$ values being in excess of 0.66 (Fig. 1; Table 1).

### Table 1. Regression coefficients of observed versus simulated mean daily flows for time series data, double mass plots and IHA analyses at four gauging weirs

| Gauging weir | $R^2$ (obs vs. sim) | $R^2$ (double mass) | $R^2$ (IHA) |
|--------------|---------------------|---------------------|-------------|
| B6H001       | 0.33                | 0.99                | 0.93        |
| K4H003       | 0.72                | 0.99                | 0.66        |
| U2H006       | 0.05                | 0.98                | 0.98        |
| X3H003       | 0.66                | 0.96                | 0.95        |

![Figure 1](https://example.com/figure1.png)

**Figure 1.** Linear regressions between IHA metrics for simulated and observed flow time-series for gauging weirs U2H006 and X3H003 in the Mgeni and Sabie River catchments respectively

Water SA 46(3) 383–392 / Jul 2020
https://doi.org/10.17159/wsa/2020.v46.i3.8648
Figure 2. Hydrological regions based on the six rainfall regions of Schulze and Maharaj (2007), and showing primary catchments A–W

Definition of hydrological regions and classification of flow types

Eight hydrological regions were defined (Fig. 2), based on agglomerations of primary catchments, WMAs and rainfall regions (Table 2). A region typically defines a spatial area of homogeneity based on a set of characteristics, and defined for a particular purpose. This provided the platform for undertaking regional redundancy analyses of flow metrics. Scree plots showed that the first two principal component axes accounted for the majority of the site variation. Many of the remaining IHA metrics showed a high degree of correlation which, after elimination based on correlations between variables and eigenvalues, provided the basis for the final PCAs based on 8–28 variables, and an average redundancy in metrics of 87%. Flow type groups were derived using PCAs in tandem with cluster classifications. In the example selected and shown here, 6 distinct flow types were defined for Flow Region 3 (late summer rainfall; middle and lower Vaal and Upper Orange River catchments). Flow types were distinguished on the basis of 28 metrics, with PC Axis 1 separating types based on flow volumes, while PC Axis 2 separated flows in terms of flow variability (Fig. 3; Table 3). Thus, flow types showed increases in median flow volume metrics from right to left, while variability increased from top to bottom.

Table 2. Level I hydrological regions for South Africa, with associated water management areas, primary catchments, and main rainfall region

| Group | WMAs                  | Primary catchment | Rainfall region |
|-------|-----------------------|-------------------|-----------------|
| 1     | Berg, Olifants/Doorn, Breede | E, G, H          | Winter          |
| 2     | Lower Orange          | D, F              | Autumn/winter   |
| 3     | Middle & Lower Vaal, Upper Orange | D             | Late summer     |
| 4     | Gouritz, Fish to Tsitsikama | J, L, N, Q     | Autumn          |
| 5     | Coastal regions of Gouritz, Fish to Tsitsikama | K, M, P, R | All year        |
| 6     | Mzimvubu, Mvoti       | S, T, U          | Mid-/late summer|
| 7     | Upper Vaal, Limpopo, Olifants, Crocodile (west), Luvuvu  | A, B, C        | Early/mid-summer|
| 8     | Inkomati, Thukela, Usuthu | V, W, X        | Mid-summer      |
Table 3. Eigenvectors for the first two axes of a PCA based on IHA metrics of flow data for quinaries from Region 3 with a late summer rainfall (WMAs = Middle & Lower Vaal, Upper Orange; Primary Catchment D). Shaded cells are for metrics reflecting coefficients of dispersion; unshaded cells are for metrics reflecting median values.

| Axes        | 1     | 2     |
|-------------|-------|-------|
| Eigenvalue  | 12.011| 10.246|
| cumul. % var.| 42.898| 79.490|
| Mean annual flow | 0.261 | −0.130 |
| 30-day maximum | 0.262 | −0.130 |
| February low flow | 0.259 | −0.130 |
| March low flow | 0.260 | −0.131 |
| May low flow | 0.259 | −0.132 |
| High flow peak | 0.262 | −0.129 |
| High flow rise rate | 0.262 | −0.127 |
| High flow fall rate | −0.262 | 0.129 |
| Small flood peak | 0.263 | −0.127 |
| Small flood rise rate | 0.254 | −0.118 |
| Large flood peak | 0.259 | −0.126 |
| Large flood rise rate | 0.244 | −0.118 |
| October | −0.123 | −0.235 |
| January | −0.119 | −0.191 |
| May | −0.095 | −0.223 |
| June | −0.110 | −0.239 |
| August | −0.124 | −0.258 |
| 30-day minimum | −0.115 | −0.251 |
| 90-day maximum | −0.102 | −0.201 |
| Baseflow index | −0.113 | −0.223 |
| Fall rate | 0.122 | 0.237 |
| October low flow | −0.112 | −0.234 |
| November low flow | −0.116 | −0.225 |
| December low flow | −0.107 | −0.212 |
| January low flow | −0.105 | −0.197 |
| February low flow | −0.098 | −0.197 |
| July low flow | −0.097 | −0.210 |
| September low flow | −0.104 | −0.243 |

Figure 3. PCA of quinary catchments in Region 3 based on the optimal metric matrix accounting for cumulative percentage variances of 42.9 and 72.5 for Axes 1–2 (see Table 3 for eigenvectors).
Flow types per hydrological region (1–8). Note that corresponding flow type numbers between hydrological regions are not equivalent: for example, a Flow Type 3 in Group 1 is not equivalent to a Flow Type 3 in Group 5.

Based on the cluster classifications, 6 to 12 flow types were defined for each of the 8 hydrological regions, with flow pattern variation explained by an optimal combination of 8–28 metrics (Fig. 4; Table 4), and yielding a total of 71 flow types nationally. While the spatial database includes all metrics for each quinary sub-catchment, the optimal metric combination for each hydrological region was different (Table 5). Cumulative percentage variation between sites was 65–89% for PC Axes 1 and 2, and an average cumulative variation of 81.25%. Axis 1 accounted for 42–71% of this variation, while Axis 2 accounted for 16–39% of the variation in flow patterns between sites. On average, metrics describing flow volumes were 1.8 times more important than metrics describing variability, although both in combination were necessary for defining flow types. In general, quinary catchments further downstream were more efficiently described by flow volume metrics, while tributary catchments and upstream catchments were more efficiently described by coefficient of dispersion metrics.

Flow types exhibited considerably more spatial heterogeneity in the higher rainfall regions of South Africa (Groups 1, 6, 8) versus those in the lower rainfall regions (Group 2). Regional differences in metrics explaining flow patterns were clearly apparent, where different combinations of metrics emerged. Each flow type represents a statistically different cluster of sub-catchments based on a combination of flow volume and dispersion metrics along a gradient of high to low values for PC Axes 1 and 2.
Table 5. Flow metrics defining flow types per hydrological region; metrics in unshaded cells are for median values, while metrics in grey shading are metrics associated with coefficients of dispersion

| Metric                                 | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
|----------------------------------------|---|---|---|---|---|---|---|---|
| Mean annual flow                       | X | X | X | X |   |   |   | X |
| Annual C. V.                           |   |   |   | X |   |   |   | X |
| Flow predictability                    | X |   | X |   | X |   |   | X |
| November                               | X |   |   |   |   |   |   |   |
| April                                  | X |   | X |   |   |   |   |   |
| 90-day minimum                         | X | X | X |   |   |   |   |   |
| 30-day maximum                         | X | X | X |   |   |   |   |   |
| 90-day maximum                         | X | X | X | X | X |   |   | X |
| Base flow index                        | X |   |   |   |   |   |   |   |
| Date of minimum                        | X |   |   |   |   |   |   |   |
| Low pulse count                        |   | X |   |   |   |   |   |   |
| High pulse count                       | X |   |   |   |   |   |   |   |
| Low pulse threshold                    | X |   | X |   |   |   |   |   |
| High pulse threshold                   | X |   | X |   |   |   |   |   |
| Rise rate                              | X | X |   | X |   |   |   | X |
| Fall rate                              | X | X |   | X | X | X | X | X |
| Number of reversals                    |   | X | X | X | X | X | X | X |
| February low flow                      | X |   |   |   |   |   |   |   |
| March low flow                         | X |   |   |   |   |   |   |   |
| May low flow                           | X |   |   |   |   |   |   |   |
| July low flow                          |   | X | X | X | X | X | X | X |
| Extreme low peak                       | X |   |   |   |   |   |   |   |
| High flow peak                         | X | X | X | X |   |   |   |   |
| High flow frequency                    | X |   |   |   |   |   |   |   |
| High flow rise rate                    | X | X | X | X | X |   |   | X |
| High flow fall rate                    | X | X | X | X | X | X | X | X |
| Small flood peak                       | X | X | X | X | X | X | X | X |
| Small flood rise rate                  | X | X | X | X | X | X | X | X |
| Small flood fall rate                  | X | X | X | X | X | X | X | X |
| Large flood peak                       | X | X | X | X | X | X | X | X |
| Large flood duration                   | X | X | X | X | X | X | X | X |
| Large flood rise rate                  | X | X | X | X | X | X | X | X |
| Large flood fall rate                  | X | X | X | X | X | X | X | X |
| October                                | X | X | X | X |   |   |   | X |
| November                               | X |   |   |   |   |   |   |   |
| January                                | X | X | X |   |   |   |   |   |
| February                               | X |   |   |   |   |   |   |   |
| April                                  | X |   | X |   |   |   |   |   |
| May                                    | X |   | X |   |   |   |   |   |
| June                                   | X |   | X |   |   |   |   |   |
| August                                 | X | X | X | X |   |   |   |   |
| September                              | X | X | X | X | X | X | X | X |
| 1-day minimum                          | X | X | X | X | X | X | X | X |
| 3-day minimum                          | X | X | X | X | X | X | X | X |
| 7-day minimum                          | X | X | X | X | X | X | X | X |
| 30-day minimum                         | X | X | X | X | X | X | X | X |
| 90-day minimum                         | X | X | X | X | X | X | X | X |
| 1-day maximum                          | X | X | X | X | X | X | X | X |
| 3-day maximum                          | X | X | X | X | X | X | X | X |
| 7-day maximum                          | X | X | X | X | X | X | X | X |
| 30-day maximum                         | X | X | X | X | X | X | X | X |
| 90-day maximum                         | X | X | X | X | X | X | X | X |
| Base flow index                        | X | X | X | X | X | X | X | X |
| High pulse threshold                   | X | X | X | X | X | X | X | X |
| Rise rate                              | X | X | X | X | X | X | X | X |
| Fall rate                              | X | X | X | X | X | X | X | X |
| October low flow                       | X | X | X | X | X | X | X | X |
| November low flow                      | X | X | X | X | X | X | X | X |
| December low flow                      | X | X | X | X | X | X | X | X |
| January low flow                       | X | X | X | X | X | X | X | X |
| February low flow                      | X | X | X | X | X | X | X | X |
| March low flow                         | X | X | X | X | X | X | X | X |
| June low flow                          | X | X | X | X | X | X | X | X |
| July low flow                          | X | X | X | X | X | X | X | X |
| August low flow                        | X | X | X | X | X | X | X | X |
| September low flow                     | X | X | X | X | X | X | X | X |
| Large flood fall rate                  | X | X | X | X | X | X | X | X |
| EFC high flow threshold                | X | X | X | X | X | X | X | X |
| EFC extreme low flow threshold         | X | X | X | X | X | X | X | X |
Examples of application

By way of illustration of how this spatial database may be used, the range of values for particular metrics plotted using box-and-whisker graphs reflects the range of variation within a selected flow type (monthly median flows across a reference hydrological year), and between flow type within a region (range of flow predictabilities; Fig. 5). For a site-specific application, we selected the appropriate quinary sub-catchment to identify the corresponding region and flow type: \( n = 67 \) for Flow Type 9 from Region 1. Flows downstream of Stettynskloof Dam for the hydrological year 2014/2015 were shown to be within reference conditions, despite being elevated above median values, for all months of the year except July (Fig. 6).

DISCUSSION

With a classification process, the key problems with using such a large number of variables relate to statistical (multi-collinearity) and logistical (time and resources in processing) challenges. Both problems are addressed by reducing the number of variables using a range of redundancy analysis techniques that included correlation matrices and PCA. However, these data can be summarized more elegantly to between 2 and 4 indices (i.e. first index for each PC axis), but also such an approach provides a framework for flow regime classification (Olden and Poff, 2003). Furthermore, results indicated that the IHA method adequately represented the majority of variation explained by 171 variables assimilated from global flow assessment literature, and provides a balance between objective selection of high-level information indices and ease of computation.

In terms of number of flow type classes, the number of our groups per hydrological region conforms to recommendations from international best practice. Our flow regions are defined at a 1:500 000 scale, and are groupings of primary catchments where broad hydrological regimes are likely to be similar based on rainfall seasonality. Cognisance is also taken of administrative practicalities, by also considering water management areas (WMAs). While the number of river types in a region should reflect the region’s heterogeneity, there remains a trade-off between detail and interpretability, with 4–12 classes recommended (Poff et al., 2010, and citing others). Thus, Kennard et al. (2010) used 120 hydrologic metrics applied to 15–30 years of flow time-series data from 830 gauging weirs across Australia to develop 12 flow type classes. Numerous approaches may be used to define classes; these have included the use of principal component analysis (Hughes and James, 1989) and cluster analysis (Poff and Ward, 1989).

Attributing a spatial database with metrics and flow types provides considerable utility value in terms of interrogating sub-catchments for baseline hydrological metrics, and rapidly defining statistically robust reference ranges for selected flow metrics. While the geographical distribution of flow types may show varying degrees of spatial cohesion, or non-contiguous distribution, an explicit spatial product resulting from regional flow type classifications could facilitate researchers developing generalisations about
interactions between hydrology and ecology, on the basis that response to natural flow regime characteristics is likely to vary between flow types (Kennard et al., 2010). The regional approach has application in smaller-scale regional planning purposes as a framework for better understanding the hydrological nature of South Africa's rivers. Beyond this, site-specific assessments would be required that take into account each river system’s unique attributes, for individual, project-level assessments (for example, individual dam or other water resource projects).

The flow type classification provides a platform for comparing flow patterns between primary catchments within the same flow region. The classification encapsulates both magnitudes and variability, and has the potential to be applied at a Sustainability Boundary Approach level (Richter, 2010), or to be prescriptive in relation to specific metrics. This product has considerable utility value through the linking of the spatial product of flow types and the associated metrics databases. Our classification could be applied in a range of arenas – including land-use change scenarios, definitions of reference flow conditions, impacts of in-channel impoundments, impacts of inter-basin transfer schemes (both donor and receiving systems), climate change impacts, and regional conservation planning and species recovery plans – and could be applied by a wide range of users and organisations required to make recommendations on flow requirements, including government departments, conservation agencies and research organisations. Scenario testing and risk analyses based on future flow scenarios can be applied to the different flow types, on the basis that these types differ in flow variability and predictability (Kennard et al., 2010).

Future research areas could include the development of scripts in the statistical software R (R Development Core Team, 2014) to automate future iterations, such as for regional refinement of the simulated data time-series. Developing biological response curves would improve the predictive power of the flow type classification through an understanding of the probable responses of key aquatic species to changes in flow metrics, where a promising approach is the use of multi-metric logistic regression models (Rivers-Moore et al., 2007; McManamay et al., 2013). This provides the basis for calculating odds ratios, thereby providing the capacity to make more informed decisions based on likelihoods, which can be taken to stakeholders for choosing the most societally acceptable regional option. Given the potential of this approach, and taking cognisance of the above points, the ultimate potential of this approach would be to work towards an online spatial database available for interrogation. Over time, and with sustained inputs, there is scope for building up a range of biological response curves that could act as hypothesis of system change that could be tested through ongoing adaptive management approaches.

ACKNOWLEDGEMENTS
The Water Research Commission of South Africa (WRC) is gratefully acknowledged for funding this research (WRC Project K5/2315). We thank Mr Wandile Nomquphu of the WRC for supporting the project. Bruce Paxton and Jarryd Gilham are thanked for their assistance in data processing.

REFERENCES
BROWN CA and JOUBERT A (2003) Using multicriteria analysis to develop environmental flow scenarios for rivers targeted for water resource management. Water SA 29 365–374. DOI 10.4314/wsa.v29i4.5041
BUNN SE and ARTHINGTON AH (2010) Basic principles and ecological consequences of altered flow regimes for aquatic biodiversity. Environ. Manage. 30 492–507. DOI 10.1007/s00267-009-9237-0
CHIEW FHS, MCMAHON TA and PEEL MC (1995) Some issues of relevance to South African streamflow hydrology. Proc. 7th S Afr. Hydrol. Symp., 4–6 September 1995, Grahamstown, South Africa.
COLWELL RK (1974) Predictability, constancy and contingency of periodic phenomena. Ecology 55 1148–1153. DOI 10.2307/1940366
DALLAS HF (2004) Spatial variability in macroinvertebrate assemblages: comparing regional and multivariate approaches for classifying reference sites in South Africa. Afr. J. Aquat. Sci. 29 161–171. DOI 10.2989/16085910409503807
DE MOOR FC (2002) Shortcomings and advantages of using rapid biological assessment techniques for the purpose of characterising rivers in South Africa. Verh. Int. Verein. Limnol. 28 651–662. DOI 10.1080/03680707.2001.11901795
DE MOOR FC (2006) Regional biogeographical differences in Trichoptera diversity in South Africa: Observed patterns and processes. In: Bueno-Soria J, Barba-Álvarez R and Armitage B (eds) Proceedings of the XIP International Symposium on Trichoptera, 18–22 June 2006. 211–218.
