Relationships between low-flow characteristics of South African streams

VY Smakhtin[1]* and M Toulouse[2]

[1]Institute for Water Research, Rhodes University, PO Box 94, Grahamstown, 6140, South Africa

[2]Ecole Nationale Supérieure Agronomique de Rennes, 35042 Rennes Cedex, France

Abstract

The low-flow regime of a river may be described in terms of various characteristics (indices). These characteristics may have multiple use in different areas of water-related research and practice and uncertainty often exists as to which low-flow characteristic is the most suitable for a specific purpose. The paper illustrates that many low-flow characteristics exhibit a strong intercorrelation and therefore one low-flow index may often be derived from another by means of regression relationships. Consequently, only one "primary" low-flow index may need to be initially estimated (either from available time series or by regionalisation techniques). The study examines various types of low-flow indices which have been estimated from more then 200 gauged daily streamflow records in South Africa. Several indices extracted from one day long-term annual flow duration curves have been used as primary characteristics. The relationships between low-flow characteristics are illustrated on the scale of the entire country and on the scale of two smaller drainage regions.

Abbreviations

ADF	Average daily flow
Q95, Q90, Q75, Q50	Daily flows extracted from flow duration curve and exceeded 95, 90, 75 and 50% of the time respectively
T_0	% of the time with zero flow conditions in a river
7Q2, 7Q10	7 d average minima with return periods of 2 and 10 years respectively
MAM7	mean annual 7 d average minimum flow
MAR	mean annual runoff
BFI	baseflow index
REC50	median ratio of daily flow recession
SD50	mean of the annual series of maximum spell durations below the threshold flow of 50% of ADF
DEF50	mean of the annual series of maximum flow deficits below the threshold flow of 50% of ADF

Introduction

Because of the diversity of water-related problems and the large variety of flow regimes, ranging from perennial rivers to ephemeral streams, low flow in South Africa is normally perceived as a dynamic concept which is not easily tied to a single characteristic or estimation method. Some sources indicate that the wide variation in low-flow characteristics in the country makes the selection of a single, predefined design flow impractical (Department of Water Affairs and Forestry, 1995) implying that low-flow assessment may be done on a case- or site-specific basis. Consequently, in many water-related areas no strict guidelines are currently established as to what low-flow characteristic to use, and the preference is given (where possible) to a representative streamflow time series from which various low-flow indices may be estimated. For the design of small water projects, water quality calculations and some ecohydrological problems, a time series of daily flows (observed or simulated) is normally a preferred option. Otherwise low-flow estimation is often performed using the regional deficient flow - duration - frequency curves (Midgley et al., 1994), which have been constructed using the simulated monthly flow time series and aim at the estimation of *n* months low-flow volumes of the specified return periods. Overall, the choice of low-flow characteristic for a specific purpose is currently not restrictive and the suitability of indices adopted elsewhere sometimes becomes a subject for separate studies (e.g. Harris and Middleton, 1993).

Smakhtin et al. (1995) calculated a variety of low-flow characteristics from daily streamflow records at about 240 gauging stations in South Africa and constructed maps of these characteristics illustrating the general pattern of their spatial distribution in the country. The indices considered for this exercise represented different aspects of low-flow regime: flows of different percentiles from the flow duration curve, flows of different return periods, recession and baseflow characteristics, etc. This study has demonstrated that low-flow regimes exhibit a high degree of spatial variability and are very dependent on local physiographic factors. At the same time, many indices demonstrate a similar spatial pattern, which implies that similar mechanisms have similar relative effects on a range of low-flow characteristics and suggested that strong correlation should exist between the indices describing different aspects of the low-flow regime. It also implies that it may be possible to establish some "primary" low-flow characteristic from which all (or most) other "secondary" low-flow indices are derived using regression models.

The choice of a primary low-flow index and the method of its estimation represent separate issues. If a representative long daily flow time series at the site of interest is available, any required low-flow index may be estimated directly from it, provided the relevant estimation software is available (Smakhtin et al., 1995). The ungauged sites pose a different problem and should be approached using the regionalisation techniques.

*To whom all correspondence should be addressed.
☎ (046) 622-4014; fax (046) 622-4014; e-mail smakhtin@iwr.ru.ac.za
Received 11 September 1997; accepted in revised form 30 January 1998

Smakhtin et al. (1997) have demonstrated a simple method by which to establish 1 d long-term annual and seasonal flow duration curves at ungauged sites. Pitman (1993) developed the disaggregation technique which may be applied to monthly flow data to calculate daily flow duration curves. Flow duration curves are widely used in the national hydrological practice and it is logical to accept, in this context, that a primary low-flow index (or indices) should be estimated from a flow duration curve.

The current study investigates the relationships between low-flow indices extracted from a 1 d long-term annual flow duration curve and several other types of low-flow characteristics representing different aspects of low-flow regime. The list of indices used in this study is given in the section below. All the indices have been calculated using low-flow estimation software which forms part of the HYMAS (HYdrological Modelling Application Software) package (Smakhtin et al., 1995). The indices have been calculated for more than 200 gauging stations with stationary records. The stations used are located upstream of all major impoundments or abstractions and have a mean record period of 20 years. In some cases only part of the record period (pre-impoundment) has been used to ensure that only non-regulated flow regimes are considered. The areas of the catchments are normally < 1 000 km^2. The details of the gauges used, the full range of calculated indices and tables containing calculated low-flow index values are presented in Smakhtin and Watkins (1997). The possibility of converting primary low-flow indices into secondary indices is investigated on a national scale (using the data from all parts of the country) and on the scale of two smaller drainage regions (one in the Eastern Cape and one in the Mpumalanga Province).

Low-flow indices: primary and secondary

The primary low-flow indices used in this study have been estimated from a 1 d long-term annual flow duration curve constructed for each selected flow gauge from available observed daily records. A flow duration curve represents the summary of the flow regime at a site displaying the complete range of recorded (simulated) discharge values from low flows to flood events. The low-flow indices that are extracted from a flow duration curve are those which are exceeded more than 50% of the time. The following three indices have been initially considered:

- **The flow which is exceeded 95% of the time (Q95)**. This index is used in some countries (e.g. the UK) for licensing of surface water abstractions. It is also widely used for the assessment of effluent discharge limits to the receiving streams.
- **The flow which is exceeded 75% of the time (Q75)**. Low-flow estimation techniques developed for some southern African countries are based on this index (Drayton et al., 1980). It is also close to the concept of "normal flow" used in South African Water Law. Smakhtin and Watkins (1997) have estimated Q75 and Q95 for a number of South African catchments from either observed or simulated time-series data.
- **The percentage of time with zero-flow conditions (T_0, %)**. Many rivers in South Africa have long zero-flow periods and this index may be important for distiguishing between perennial, intermittent and ephemeral streams. T_0 may be read directly from the flow duration curve graph at a point where the curve intersects the time axis.

The emphasis has been placed on the first two indices. T_0 has been used mostly to differentiate between perennial and non-perennial streams to specify subsets of data for regression analysis. It has also been used in several regression models as the second independent variable. Dependent on the type of the secondary index being considered for a relationship, Q95 and Q75 have been expressed either in m^3/s or as a proportion of long-term ADF.

The following secondary low-flow indices have been considered:

- Discharges estimated from the annual series of flow minima. They include the 7 d average flow with return periods of 10 and 2 years (7Q10 and 7Q2) and the 7 d average MAM7. The first two are used as designed low flows in the USA, while the latter is the alternative index used in the UK for abstraction licensing.
- BFI. The BFI represents the general baseflow response of a catchment, frequently used to study the effects of catchment geology on low flows, and is estimated as the volume of baseflow divided by the volume of total streamflow (the volume of baseflow may be calculated by digital filtering from continuous daily streamflow data).
- REC50. This characteristic represents the rate of baseflow recession and is estimated from the distribution of daily recession ratios (today's flow divided by yesterday's flow) calculated for all recession periods found in a record for those days when discharge is less than ADF. This index is similar to the one described in FREND (1989). It may reflect the effects of catchment geology on low flows or serve as a criterion for evaluating rainfall-runoff daily model ability to simulate low flows (Smakhtin et al., 1998).
- Characteristics of continuous low-flow events (spells or runs). A low-flow spell is defined as an event when the flow is continuously below a certain specified threshold discharge. Each low-flow spell is characterised by the duration and the deficit which would be required to maintain the flow at a given threshold. Spell analysis normally deals with the annual time series of maximum durations/deficits, from which a number of spell indices may be extracted (Smakhtin and Watkins, 1997). The indices used in this study include the mean duration and deficit of spell maxima below 50% of ADF (further referred to as SD50 and DEF50 correspondingly). SD50 has been expressed as the proportion of the number of days in a year (SD50/365), while DEF50 is expressed as a percentage of the MAR. The duration and deficit of low-flow spells with return period of 10 years were also considered initially, but it was found that their relationships with primary low-flow indices are very similar to those of SD50 and DEF50. Therefore, these results are not specifically reported here.

More detailed descriptions of various low-flow indices (including all those used in the current study) and the techniques for their estimation may be found in Smakhtin et al. (1995) or Smakhtin and Watkins (1997).

For each secondary low-flow index a number of various regression model types have been examined. They include:

- Linear model;
- Logarithmic model 1 (both dependent and independent variables are transformed into natural logarithms);
- Logarithmic model 2 (dependent variable only transformed into natural logarithm);

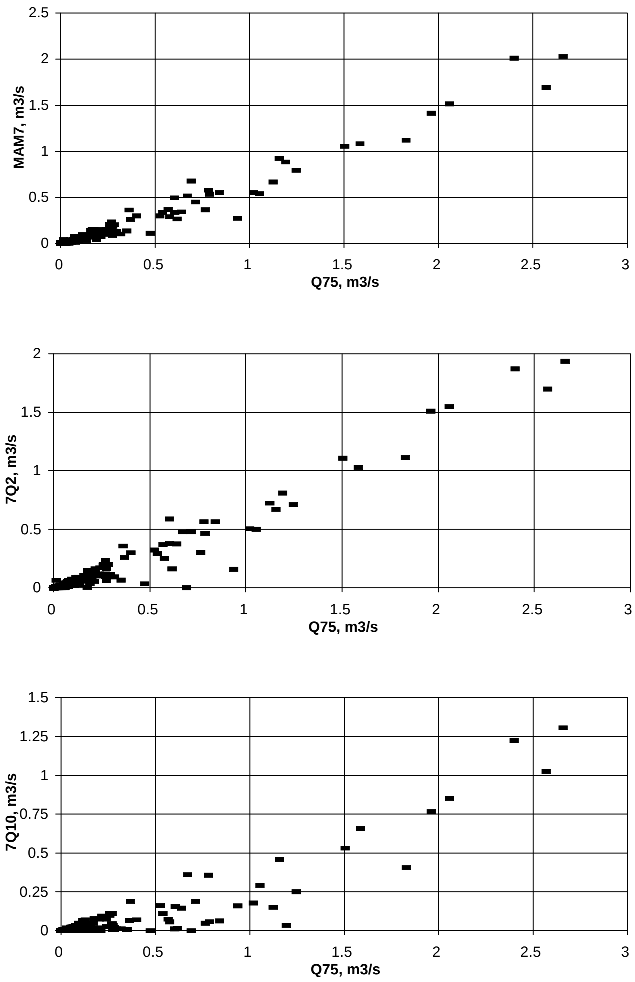

Figure 1

Scatter plots of the relationship between Q75 and low-flows of different frequency of occurrence

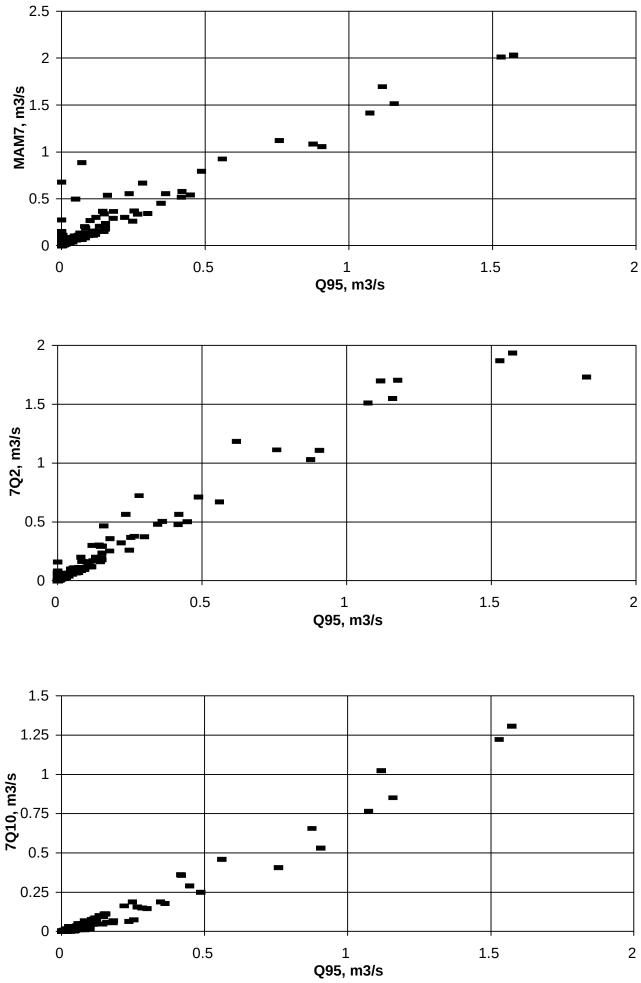

Figure 2

Scatter plots of the relationship between Q95 and low-flows of different frequency of occurrence

- Square root model 1 (both dependent and independent variables are transformed);
- Square root model 2 (only dependent variable is transformed);
- Inverse model (for dependent variable);

This paper presents only the best (in terms of statistical criteria) models for each secondary low-flow characteristic.

Results and discussion

The scatter plots of the relationship between primary indices extracted from the flow duration curve and other low-flow characteristics are shown in Figs. 1 to 4. Table 1 lists the best regression models established with corresponding values of coefficient of determination (R^2) and standard error (SE). The sample size (the number of gauges) used in most of the cases was 208.

Figure 1 demonstrates a strong relationship between Q75 flow and all three low-flow characteristics estimated from the series of annual minima (MAM7, 7Q10 and 7Q2). Q75 is especially strongly correlated with MAM7 and 7Q2. Estimates for several perennial rivers have shown that MAM7, if placed on the 1 d long-term annual flow duration curve, is exceeded 80 to 91% of the time, and 7Q2 is exceeded 83 to 93% of the time. The values of MAM7 and 7Q2 for each particular stream are therefore generally close to each other. Both are also closer to Q75 than to Q95 and show a slightly worse correlation with the latter (Fig. 2 and Table 1).

7Q10 is an index of more extreme low-flow conditions. On the flow duration curve it is normally exceeded 95 to 99.5% of the time (in the majority of cases tested it was exceeded more than 99% of the time). It therefore correlates better with Q95 than with Q75 (Fig. 2 and Table 1). Consequently, MAM7 and 7Q2 show better relationships with Q75, while Q95 is a better predictor for 7Q10.

The relationship between baseflow and recession indices (BFI and REC50) and Q75 is illustrated by the scatter plots in Fig. 3 (similar graphs are obtained for Q95). Unlike the “frequency” indices, these two ratios are not the actual flows and, although estimated from the streamflow time-series data, they rather represent the generalised characteristics of the subsurface storage of a catchment. Consequently, their relationships with particular low-flow values are not that explicit. The characteristic

TABLE 1 RELATIONSHIPS BETWEEN LOW-FLOW INDICES				
Low-flow index	**Regression model**	**R^2**	**SE**	**Sample size**
MAM7, m^3/s	MAM7 = 0.691*Q75 MAM7 = 1.37*Q95	0.97 0.91	0.06 0.10	208 208
7Q2, m^3/s	7Q2 = 0.658*Q75 7Q2 = 1.33*Q95	0.94 0.93	0.07 0.08	208 208
7Q10, m^3/s	7Q10=0.343*Q75 7Q10=0.744*Q95	0.82 0.97	0.08 0.03	208 208
BFI	BFI =0.229 + 0.62*(Q75/ADF) -0.002*T_0 BFI=0.279+ 0.856*(Q95/ADF) - 0.003*T_0 BFI = 0.20 + 0.712* (Q75/ADF)	0.78 0.72 0.73	0.07 0.07 0.06	208 208 158
REC50	REC50 =0.957 - 0.0034*T_0	0.60	0.07	208
SD50/365	ln SD50 = -0.94 - 1.74*(Q75/ADF) ln SD50 = -1.04 -2.70*(Q95/ADF)	0.43 0.33	0.27 0.29	208 208
DEF50, %MAR	ln DEF50 = 2.88 -3.94*(Q75/ADF) ln DEF50 = 2.58 -5.8*(Q95/ADF) +0.005*T_0 ln DEF50 = 2.9 -4.0*(Q75/ADF)	0.79 0.72 0.79	0.27 0.32 0.26	208 208 158

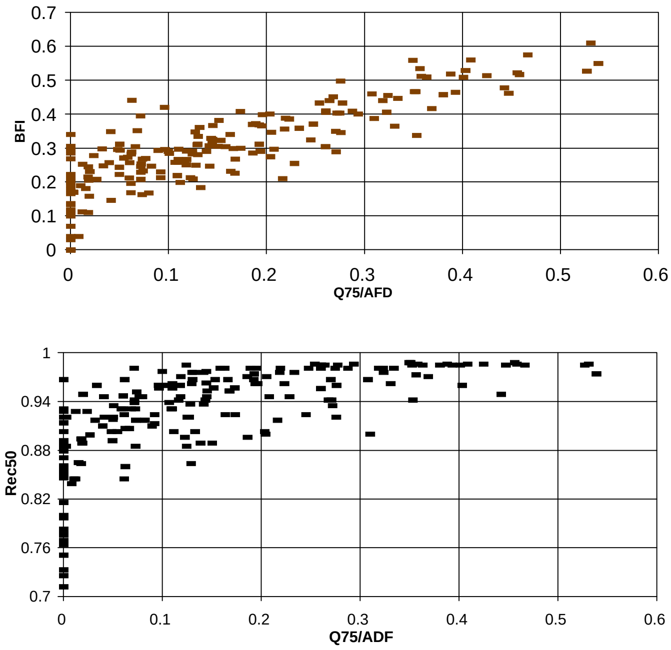

Figure 3
Scatter plots of the relationship between Q75, baseflow index (BFI) and median recession ratio (REC50)

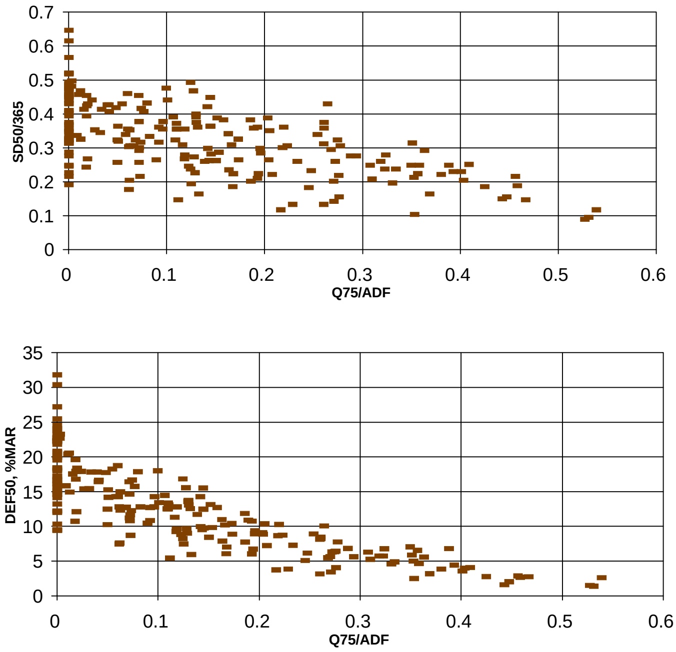

Figure 4
Scatter plots of the relationship between Q75 and characteristics of continuous low-flow events below the threshold discharge of 50% ADF

TABLE 2 RELATIONSHIPS BETWEEN LOW-FLOW INDICES IN X DRAINAGE REGION			
Low-flow index	**Regression model**	**R^2**	**SE**
MAM7, m^3/s	MAM7 = 0.647*Q75	0.99	0.09
7Q2, m^3/s	7Q2 = 0.727*Q75	0.99	0.08
7Q10, m^3/s	$7Q10^{0.5}$ =0.033 +0.343*$Q95^{0.5}$	0.87	0.12
BFI	BFI =0.272 + 0.583*(Q75/ADF) -0.003*T_0	0.84	0.03
REC50	$REC50^{0.5}$ =0.962 - 0.0016*T_0+ 0.07*(Q75/ADF)	0.73	0.01
SD50/365	SD50 =0.437 - 0.56*(Q75/ADF)- 0.005*T_0	0.72	0.04
DEF50, %MAR	DEF50 = 14.6 -25.6*(Q75/ADF)	0.89	1.05

feature of the scatter plots is that a zero Q75 flow value corresponds to a large range of non-zero BFI and REC50 values.

In the first case, this is partially related to the limitations of the digital filtering technique used to calculate BFI from time-series data. Smakhtin et al. (1995) have found that digital filtering has a tendency to overestimate BFI especially for intermittent streams as it often creates excessive baseflow for isolated relatively short-term flood events. For the purpose of this study, the intermittent streams may be defined as those having zero Q75 (and therefore, T_0 not less then 25%). However, this definition is rather arbitrary since baseflow overestimation by the filter could have also occurred (although to a lesser extent) in cases when T_0 was less than 25%.

If only Q75 (or Q95) is used as a predictor in the regression model for BFI, the coefficient of determination does not exceed 0.65. In order to improve the regression relationship, all intermittent rivers were excluded from the data set. In this case, the sample size was reduced to 158, but R^2 increased to 0.73. Alternatively, T_0 may be included in regression model as a second independent variable. The results achieved using both options are only marginally different (Table 1). The relationships between BFI and Q75 are slightly better than that those between BFI and Q95 possibly because Q75 (as well as BFI itself) is a better indicator of average baseflow response of a catchment than the more extreme Q95.

In the case of REC50, 60% of its variability is explained by T_0 (Table 1), whereas the R^2 of the linear regression between REC50 and either Q75 or Q95 does not exceed 0.26. If all intermittent rivers (with T_0 more then 25%) are excluded from the data set, the R^2 of each REC50 individual relationship with the two primary low-flow indices is in the range of 0.32 to 0.42.

The major problem, which may also partially explain the lack of correlation between REC50 and other indices, possibly relates to the method of estimation of REC50 itself from the time series data. REC50 should represent the rate of baseflow recession. In the estimation procedure, the recession ratios were calculated for each and every day of recession when discharge is less then ADF. For many streams, the minor flow fluctuations which occur on top of the main recession limb (formed by surface or quick subsurface flow), may effect the results of the estimation process and cause the underestimation of the final REC50 value. It may therefore be necessary to adjust the estimation technique of REC50 and to recalculate it for every stream using only continuous (8 to 10 d and longer) recession limbs unaffected by minor flow fluctuations. This may result in REC50 values which are more representative of actual baseflow recession in each stream. Also the analysis should only be done for streams with well defined recession properties and therefore only perennial streams should be included. On the other hand, it is possible that for establishing a better regression model, other primary indices will have more value. For example, the Q90/Q50 ratio, which characterises the slope of the entire lower portion of the flow duration curve, may be a better predictor than any particular low-flow discharge.

The scatter plots of spell characteristics SD50 and DEF50 with Q75 are shown in Fig. 4 (similar relationships have been obtained between spell characteristics and Q95). Table 1 illustrates that spell duration can hardly be predicted from the selected primary low-flow indices with reasonable accuracy. This is due to the fact that a low-flow spell may be interrupted by minor increases in flow (dry season freshes) and therefore the duration of low-flow spells is more dependent on variability of daily flows rather than on a single low-flow characteristic. This relationship may be investigated in more detail using seasonal flow duration curves and/or using characteristics of daily flow variability.

Good relationships have been established between primary low-flow indices and spell deficits (Table 1). This may be primarily explained by the fact that low-flow discharge values (represented by either Q75 or Q95) determine the magnitude of flow deviation from the specified threshold value (50% of ADF in this case) during a low-flow spell. Marginal improvement to the regression model was achieved when the exercise was repeated for a reduced set of perennial rivers (Table 1).

All results illustrated above have been obtained at the scale of the entire country. The data set used included streams with different types of low-flow regime. Better regression models may be established at the scale of smaller geographical regions where the physiographic conditions vary less significantly and which are consequently more hydrologically homogeneous (e.g. 22 primary drainage regions of South Africa (Midgley et al., 1994)). Tables 2 and 3 list some of the regression models which have been established in the X drainage region (in the Mpumalanga Province; sample size 23) and in drainage regions S and T taken together (in the Eastern Cape Province; sample size 15). Compared to the results obtained using the data set for the entire country, an improvement has been achieved in the prediction of BFI, REC50, and spell characteristics. One possible explanation is that the majority of rivers in these regions (whose data were used for analysis) are perennial and have well-defined recession characteristics. Consequently, the values of secondary indices are more representative and less error-prone.

TABLE 3
RELATIONSHIPS BETWEEN LOW-FLOW INDICES IN S AND T DRAINAGE REGIONS

Low-flow index	Regression model	R^2	SE
MAM7, m^3/s	MAM7 = 0.037+1.28*Q95	0.98	0.07
7Q2, m^3/s	7Q2 = 1.329*Q95	0.99	0.06
7Q10, m^3/s	7Q10 =0.706*Q95	0.98	0.04
BFI	BFI =0.169 + 0.875*(Q75/ADF) -0.003*T_0	0.93	0.04
REC50	REC50 =0.957 - 0.005*T_0+ 0.08*(Q75/ADF)	0.96	0.02
SD50/365	$SD50^{-1}$ =0.007 + 0.0028*(Q95/ADF)	0.60	0.001
DEF50, %MAR	ln DEF50 = 2.72 -5.73*(Q95/ADF) + 0.002*T_0	0.90	0.14

Conclusions

Regression relationships have been established between low-flow indices estimated from a long-term annual flow-duration curve and several other types of low-flow characteristics, representing frequency of extreme low-flow discharges, recession rate, relative baseflow contribution to streamflow and continuous low-flow periods. These relationships are derived using the data from more than 200 rivers from the whole of South Africa, and smaller data sets for two primary drainage regions in the country. The relationships are presented in Tables 1, 2, and 3 and many of them may be recommended for quick conversion of one type of low-flow characteristics into others. However, additional research is necessary to improve the regression models for characteristics of streamflow recession and duration of continuous low-flow events.

These relationships may be particularly appropriate for use at ungauged sites. A flow duration curve (as a “source” of primary low-flow indices) at an ungauged site may be established either by regionalisation of observed daily flow time series (Smakhtin et al., 1997) or by dissagregation of monthly flow data (Pitman, 1993). These techniques allow flow duration curves to be estimated without the actual observed flow time series at a site. Alternatively, the required primary low-flow indices Q75 or Q95 may be mapped on a regional basis (using observed daily data), or estimated from monthly flow data using the regression technique. Smakhtin and Watkins (1997) illustrated this approach using the streamflow gauges in the Tugela River catchment as an example. Similarly, some other low-flow characteristics (e.g. indices of different frequency of occurrence) may be derived from monthly flow data directly. Overall, the relationships between monthly flow data and various low-flow characteristics which are currently estimated on the basis of observed daily time series, may be a subject for a separate study. If the outcome of such a study is positive, it would add value to the extensive research work that has already been accomplished at the scale of the whole country (Midgley et al., 1994)

Acknowledgements

The financial support of the Water Research Commission is greatly appreciated. The daily streamflow data have been provided by the Department of Water Affairs and Forestry.

References

DRAYTON RS, KIDD CHR, MANDEVILLE AN and MILLER JB (1980) A Regional Analysis of Floods and Low Flows in Malawi. Inst. of Hydrol., Report No **72**, Wallingford, UK. 79 pp.

DEPARTMENT OF WATER AFFAIRS AND FORESTRY (1995) Procedures to Assess Effluent Discharge Impacts. Pretoria, South Africa. 352 pp.

FREND (Flow Regimes From Experimental and Network Data) (1989) *I: Hydrological Studies.* Institute of Hydrol., Wallingford, UK. 344 pp.

HARRIS J and MIDDLETON BJ (1993) Calculation of low flows for selected South African rivers and implications for water quality. *Proc. of the 6th S. Afr. Natl. Hydrol. Symp.* Vol **2** 655-664.

MIDGLEY DC, PITMAN WV and MIDDLETON BJ (1994) Surface Water Resources of South Africa 1990. WRC Report 298/5.1/94. Pretoria, South Africa.

PITMAN WV (1993) Simulation of run-of-river schemes using monthly data. *Proc. of the 6th S. Afr. Natl. Hydrol. Symp.* Vol **1** 445-452.

SMAKHTIN VY, WATKINS DA and HUGHES DA (1995) Preliminary analysis of low-flow characteristics of South African rivers. *Water SA* **21** 201-210.

SMAKHTIN VY, CREUSE-NAUDIN E and HUGHES DA (1997) Regionalisation of daily flow characteristics in part of the Eastern Cape, South Africa. *Hydrol. Sci. J.* **42** 919 - 936.

SMAKHTIN VY, SAMI K and HUGHES DA (1998) Evaluating the performance of a deterministic daily rainfall-runoff model in a low-flow context. *Hydrol. Processes.*

SMAKHTIN VY and WATKINS DA (1997) Low-flow Estimation in South Africa. WRC Report 494/1/97. Pretoria, South Africa.