Generating natural daily flow sequences for South African rivers from historical flow data

VLADIMIR SMAKHTIN, DENIS HUGHES
Institute for Water Research, Rhodes University, PO Box 94, Grahamstown 6140, South Africa

OLGA SMAKHTINA
Department of Geography, University of Fort Hare, Private Bag X1314, Alice 5700, South Africa
e-mail: olga.s@ufhcc.ufh.ac.za

Abstract Daily streamflow time-series data are often in demand in South Africa in water resource assessment problems and in river ecology studies. The latter normally require daily time-series data for natural conditions in a catchment. The generation of such time series by means of deterministic physically based daily models is a time consuming approach, while simpler methods may have a similar success. The paper investigates the possibilities of reproducing the natural daily flow variability using historical daily flow time series, which may be partially affected by water resource development. The natural daily time series at a site of interest is generated using a non-linear spatial interpolation technique, which is based on 1-day flow duration curves for each calendar month of the year. The general purpose of this technique is to transfer the daily flow information from a source site(s) with data to the destination site(s) where these data are either insufficient, not available or need to be naturalized. The paper describes the approaches by which to establish 1-day flow duration curves representing the natural flow regime at the destination site and illustrates the application of the time-series generation technique with examples from regulated streams in South Africa.

INTRODUCTION

Daily streamflow information is required in the design of run-of-river schemes, water quality calculations, assessment of catchment development impacts on river ecology etc. With the recent shift towards the integrated management of water resources, rapid development of small-scale water supply schemes in previously underdeveloped regions and increased attention to riverine ecology, these data are required urgently. At the same time the characterization of daily flow regimes from observed data in many regions of South Africa is possible only at a limited number of sites. Also, streamflow records are often affected by various anthropogenic impacts, while it is the data representing natural catchment conditions which are in demand. The typical example of the latter in South Africa is the estimation of Instream Flow Requirements (IFR) necessary to preserve the ecological balance of a stream after the implementation of the water scheme. The IFR intend to “design” a modified flow regime which would mimic the most important features of the natural one.

The use of deterministic rainfall–runoff modelling techniques is one possible and rather common solution to the problem. However, certain complications arise when
the simulation of streamflow time series in natural conditions is intended. A model has first to be calibrated against existing flow data (normally affected by catchment changes), and then its parameters have to be modified to represent the natural state of the catchment. The direction and degree of parameter changes should be scientifically justified. Eventually the model is run with the new parameter values to generate the “naturalized” flow time series (all direct abstractions/effluents to the stream are obviously ignored). Generally, the rainfall–runoff modelling is a complex and labour intensive approach, and simpler estimation techniques in certain circumstances may have similar value.

Hughes & Smakhtin (1996) described the pragmatic technique of spatial interpolation of observed flow data which was initially designed to patch/extend daily flow records at the site of interest. The technique allows a time series of daily flows from one site to be transferred to another using a flow duration curve (FDC) as a key “transfer function”. (A FDC is a relationship between a flow value and the percentage of time this flow is equalled or exceeded, and it therefore gives a “summary” of the hydrological regime at a site displaying the complete range of river discharges from low flows to flood events). This method was extensively tested using the data from different locations in southern Africa and was found to perform at least as well as a more complex rainfall–runoff simulation approach. Smakhtin et al. (1997) described a method by which the application of the spatial interpolation technique for daily flow time series generation may be extended to completely ungauged sites. The current paper investigates one of the other possible extensions of spatial interpolation technique—its application for the generation of daily flow time series representing natural conditions in a catchment.

NONLINEAR SPATIAL INTERPOLATION TECHNIQUE

The nonlinear spatial interpolation technique of observed streamflow records is based on 1-day FDCs for each calendar month of the year. The flow record at the site of interest (destination site) may either be deficient in some way (e.g. needs to be patched or extended), or completely missing. In the former case, the required set of 1-day monthly FDCs at the destination site may be established directly from the record (Hughes & Smakhtin, 1996). In the latter, FDCs may be established using regionalization techniques (Smakhtin et al., 1997). In either case a good quality flow record at least at one of the neighbouring (source) sites should be available for data transfer and the application of the spatial interpolation technique follows several subsequent steps:

- identify up to five possible “source” flow gauging stations and assign weights to each of them based on the degree of similarity between “source” and destination site flow regimes;
- generate tables of discharge values for each site and each month of the year for fixed percentage points on the FDCs;
- identify the % point of each day’s flow at the source site; read off the flow value for the equivalent % point from the destination site’s FDC table; log-interpolate between fixed % points; repeat the procedure for each “source” site and take the weighted average of all estimates for the “destination” site.
The technique has been set up in the form of a "model" which may be applied to a number of gauges in a catchment at once and "calibrated" in a similar way to, for example, rainfall-runoff models by (a) changing the selection of the source sites and their number and/or (b) changing the weighting factors for each "source" site. The simulated daily flow time series may be visually compared with original observed data (if it exists) during the calibration period. The results of simulations may also be assessed using standard criteria of fit between observed and simulated time series. The comparison is made for untransformed and log-transformed flows. The fit statistics used are the maximum, minimum and mean flow value, standard deviation and coefficients of determination ($R^2$) and efficiency (CE). $R^2$ and CE are defined as follows:

$$R^2 = \frac{\sum (S_i - MO)^2}{\sum (O_i - MO)^2} \quad (1)$$

$$CE = 1 - \frac{\sum (O_i - S_i)^2}{\sum (O_i - MO)^2} \quad (2)$$

where $O_i$ and $S_i$ are correspondingly observed and simulated values and $MO$ is the mean value of the observed time series. The purpose of comparison of untransformed flows is to assess the general quality of simulations, while fit statistics for log-transformed flows provide a better indication of the correspondence for low flows. The model output consists of (a) patched/extended observed flow time series; (b) generated flow time series which are made up completely of simulated daily flow values regardless of whether the original observed flow was missing or not.

**APPLICATION FOR SITES LOCATED DOWNSTREAM OF RESERVOIRS**

For ecological purposes it is often required to know the degree of modification of the natural flow regime in a river due to upstream flow regulation. Therefore, the representative time series (often including particular years, e.g. extremely dry) for unmodified flow conditions downstream of existing reservoirs, are required.

If daily flow records exist at the destination site located downstream of the reservoir, and the observation period at this site is only partially affected by the upstream impoundment (normally the latest part of the period), then part of the record representing unmodified flow regime at the destination site should be identified. The observations at the source gauge(s) should certainly extend into the period for which the flow generation at the destination site is intended.

A number of examples which illustrate this situation may be found in South Africa. One example is the Blyde River catchment located to the west of the Kruger National Park boarder. The gauge B6H005 (catchment area 2204 km$^2$) with a record starting in 1958, is located downstream of the Blydepoort Dam which was constructed in 1974 (Fig. 1). No catchment or water resource developments occurred upstream of this gauge until the construction of the reservoir. Therefore, the first part of the record at gauge B6H005 (prior to 1974) represents the natural flow regime in the river, while the second part (from 1974 until present) reflects the modified flow. Tables of discharge values for 17 fixed percentage points for each calendar month have therefore been generated using the unmodified flow period (1958–1973).
The suitable neighbouring source sites are B6H001 (518 km$^2$), B6H003 (93 km$^2$) and B6H006 (43 km$^2$, Fig. 1). The first gives the largest contribution to the flow at gauge B6H005 downstream. It is also the closest gauge to the destination site and is located on the same river. Therefore, it should be assigned the largest weight (e.g. 0.8). The other two contribute less and may be assigned smaller weights (e.g. 0.1 each). All source gauges record natural flow conditions and their entire observation periods (normally from the late 1950s until the present) have been used to generate discharge tables.

The statistics of fit between observed and generated flows are summarized in Table 1. High values of $R^2$ and CE for both untransformed and log-transformed flows during pre-impoundment period are indicators of a good simulation. During this period the flows in both wet and dry years are simulated equally well (Fig. 2). Relative deterioration of the fit statistics (particularly for log-transformed flows) occurs in the post-impoundment period (Table 1). This is a direct consequence of comparing observed regulated and simulated natural daily flow sequences. The differences between the two are especially pronounced during dry years, when flow regulation becomes critical (Fig. 3).

If no flow observations exist downstream of the reservoir site during the pre-impoundment period, the set of 1-day monthly FDCs for the destination site may be

### Table 1

<table>
<thead>
<tr>
<th>Period</th>
<th>Time series</th>
<th>Untransformed flows:</th>
<th>Log-transformed flows:</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Max ($m^3 s^{-1}$)</td>
<td>Mean ($m^3 s^{-1}$)</td>
</tr>
<tr>
<td>1958–1973</td>
<td>Obs.</td>
<td>318</td>
<td>6.72</td>
</tr>
<tr>
<td></td>
<td>Sim.</td>
<td>296</td>
<td>6.47</td>
</tr>
<tr>
<td>1974–1996</td>
<td>Obs.</td>
<td>270</td>
<td>7.37</td>
</tr>
<tr>
<td></td>
<td>Sim.</td>
<td>284</td>
<td>8.09</td>
</tr>
</tbody>
</table>
established using either regionalization techniques on the basis of several natural records from neighbouring gauges (the procedure is similar to that described by Smakhtin et al., 1997). The more pragmatic alternative is to use just one nearest upstream gauge with good quality data. In this case each of the fixed percentage points on the discharge table generated from these data should be adjusted accordingly using a certain correction factor to produce the discharge table at the destination site.

The correction factor may be calculated as the ratio between the catchment areas at the destination site and the site with data. However, one of the critical aspects of this approach is that streamflows at even closely adjacent sites are rarely linearly related to catchment area. The more valid alternative is to use the ratio of the Mean Annual Runoff (MAR) values of the destination site of interest and the nearest
upstream gauged site with data. In South Africa the estimates of the “virgin” MARs are available for almost 2000 small and normally ungauged drainage subdivisions throughout the entire country (quaternary subcatchments) from the report “Surface Water Resources of South Africa 1990” (Midgley et al., 1994). For larger basins, MAR may be calculated as a sum of MARs of all quaternary subcatchments above the catchment outlet. If the virgin MAR data are not readily available (e.g. for other countries), these estimates may be obtained by means of regional regression models.

Generation of natural flow sequences for a regulated river reach in the absence of historical flow records downstream of the reservoir site may be illustrated by the example of the Bushmans River located in KwaZulu-Natal Province to the east of the border with Lesotho (Fig. 4). Good quality flow records at the two adjacent gauges V7H012 (196 km²) and V7H017 (276 km²), starting normally in the early 1970s, reflect the unmodified flow conditions. The flow is also measured at the Wagendrift Dam (gauge V7H020, upstream catchment area 744 km²) constructed in 1963, but no flow records exist for downstream of the dam during the pre-impoundment period.

The flow at gauge V7H017 contributes the most to the inflow to the dam and its FDC discharge tables may be used to generate FDC discharge tables at the reservoir site (it should also be assigned the largest weight for the application of the nonlinear spatial interpolation technique). The historical MAR estimate at V7H017 is $118 \times 10^6$ m³, while the virgin MAR at the dam site (as the sum of three upstream quaternary subcatchments (Fig. 4) is $222 \times 10^6$ m³. Therefore, to calculate the FDC discharge tables at the dam site, the FDC discharge tables at V7H017 are corrected by the factor of 1.882 ($\frac{\text{MAR}_{\text{DAM}}}{\text{MAR}_{\text{V7H017}}}$). Once the FDC discharge tables at the destination dam site are constructed, the spatial interpolation technique is applied in the same way as described above. The restored natural daily streamflow time series at the dam site reflects the main features of the unmodified upstream flow regime.

Fig. 4 Map of the Bushmans River catchment showing the location of streamflow gauges and quaternary catchment boundaries.
and, if compared with observed flow at the dam, clearly illustrates the degree of changes in flow regime brought to the entire downstream river system by flow regulation (Fig. 5).

All the calculations and analyses described, including the construction of the FDCs, generation and correction of the FDC’s discharge tables, setting up and running the spatial interpolation algorithm and calculating fit statistics have been performed using the comprehensive PC based system HYMAS (Hydrological Modelling Application Software) which was developed in the Institute for Water Research of Rhodes University and is designed to run various hydrological models and analyse observed and/or simulated hydrological variables.

CONCLUSIONS

The method described in this paper is aimed at the quick restoration of natural daily streamflow sequences frequently required in ecological studies. It is designed as a pragmatic alternative to the application of more complex deterministic daily rainfall–runoff models. Although the performance of the method is illustrated with examples of regulated streams only, its application is not limited to such cases and may be extended to other situations, e.g. in which natural daily flow time series data are required for ungauged sites. As such, the technique has been widely used by the authors in recent years for various water resource assessment problems in South Africa.

Weighting of the FDC discharge table by a constant correction factor implies that both high and low flows are changing downstream to the same proportion. This may not be the case in many rivers where the natural variability of daily flows is decreasing downstream. Consequently, more research is required to address and properly reflect in the existing technique the issue of different changes in high and low flows with the increasing size of the river catchment.
Acknowledgements The study is part of a project funded by the Water Research Commission of South Africa. Thanks are due to the South African Department of Water Affairs and Forestry for the provision of streamflow data.

REFERENCES

