This article was downloaded by: [b-on: Biblioteca do conhecimento online UTL] On: 18 May 2012, At: 11:05 Publisher: Taylor & Francis Informa Ltd Registered in England and Wales Registered Number: 1072954 Registered office: Mortimer House, 37-41 Mortimer Street, London W1T 3JH, UK



Hydrological Sciences Journal

Publication details, including instructions for authors and subscription information: <u>http://www.tandfonline.com/loi/thsj20</u>

Disaggregation modelling of monthly streamflows using a new approach of the method of fragments

A.T. Silva^a & M.M. Portela^a

^a Department of Civil Engineering, Instituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal

Available online: 15 May 2012

To cite this article: A.T. Silva & M.M. Portela (2012): Disaggregation modelling of monthly streamflows using a new approach of the method of fragments, Hydrological Sciences Journal, DOI:10.1080/02626667.2012.686695

To link to this article: <u>http://dx.doi.org/10.1080/02626667.2012.686695</u>

First

PLEASE SCROLL DOWN FOR ARTICLE

Full terms and conditions of use: <u>http://www.tandfonline.com/page/terms-and-conditions</u>

This article may be used for research, teaching, and private study purposes. Any substantial or systematic reproduction, redistribution, reselling, loan, sub-licensing, systematic supply, or distribution in any form to anyone is expressly forbidden.

The publisher does not give any warranty express or implied or make any representation that the contents will be complete or accurate or up to date. The accuracy of any instructions, formulae, and drug doses should be independently verified with primary sources. The publisher shall not be liable for any loss, actions, claims, proceedings, demand, or costs or damages whatsoever or howsoever caused arising directly or indirectly in connection with or arising out of the use of this material.

Disaggregation modelling of monthly streamflows using a new approach of the method of fragments

A.T. Silva and M.M. Portela

Department of Civil Engineering, Instituto Superior Técnico, Technical University of Lisbon, Lisbon, Portugal artur.tiago.silva@ist.utl.pt; mps@civil.ist.utl.pt

Received 21 February 2011; accepted 14 October 2012; open for discussion until 1 January 2013

Editor D. Koutsoyiannis; Associate editor C. Onof

Citation Silva, A.T. and Portela, M.M., 2012. Disaggregation modelling of monthly streamflows using a new approach of the method of fragments. *Hydrological Sciences Journal*, 57 (5), 1–14.

Abstract The method of fragments is applied to the generation of synthetic monthly streamflow series using streamflow data from 34 gauging stations in mainland Portugal. A generation model based on the random sampling of the log-Pearson Type III distribution was applied to each sample to generate 1200 synthetic series of annual streamflow with an equal length to that of the sample. The synthetic annual streamflow series were then disaggregated into monthly streamflows using the method of fragments, by three approaches that differed in terms of the establishment of classes and the selection of fragments. The results of the application of such approaches were compared in terms of the capacity of the method to preserve the main monthly statistical parameters of the historical samples.

Key words generation of synthetic series; disaggregation modelling; method of fragments; definition of the classes of fragments

Modèle de désagrégation des débits mensuels utilisant une nouvelle approche de la méthode des fragments

Résume La méthode des fragments est appliquée à la génération de séries synthétiques de débits mensuels en utilisant les données de 34 stations de jaugeage au Portugal continental. Le modèle de génération basé sur l'échantillonnage aléatoire de la loi log-Pearson III a été appliqué à chaque échantillon, pour générer 1200 séries synthétiques de débits annuels de même longueur que l'échantillon de base. Les débits annuels synthétiques ont ensuite été désagrégés en débits mensuels par la méthode des fragments, suivant trois approches différentes concernant la répartition en classes et la sélection des fragments. Les résultats de l'application de ces approches ont été comparés en termes de capacité de la méthode à préserver les principaux paramètres statistiques mensuels des échantillons historiques.

Mots clefs génération de séries synthétiques; modèle de désagrégation; méthode des fragments; définition des classes de fragments

1 INTRODUCTION

Stochastic generation of hydrological data is widely applied in the planning and management of water resources systems (Salas *et al.* 1980, p. 16, Santos 1983, p. 1) The modelling of monthly hydrological data, in general, and monthly streamflow data, in particular, can adopt two general approaches: (i) the generation of monthly flows directly using an appropriate monthly model; or (ii) the generation of annual flows using an appropriate annual model, and the subsequent disaggregation of those annual flows into monthly flows. The second approach has the advantage of allowing the preservation of the statistical properties of a historical sample at both the annual and monthly time scales (Santos and Salas 1992). The method of fragments, proposed by Svanidze (1964), is one of the disaggregation models presented by the literature on stochastic streamflow generation (Srikanthan and McMahon 1980a, 1980b, 1982, Svanidze 1980, Santos 1983, Maheepala and Perera 1996, McMahon and Adeloye 2005, Ndiritu 2011). The method of fragments uses a purely deterministic approach to disaggregate annual flows into monthly flows: given a historical streamflow sample of N years, the observed monthly flows in each year of the historical sample are divided by the corresponding annual streamflow. The resulting set of 12 standardized monthly flows constitutes a fragment. Each value of the generated annual streamflow series is disaggregated into 12 generated monthly flows by a fragment. According to Santos (1983, pp. 19–20), the way in which the fragments are assigned to each generated annual value is the most important feature

of the method. After the fragments are obtained, they can be assigned randomly to the generated annual flows. Srikanthan and McMahon (1980a, 1980b, 1982) found that this approach did not preserve the monthly parameters of the historical sample; hence, they attempted to improve the method by introducing a relationship between the fragments and the annual flows. This was done by sorting the N historical annual flows from smallest to largest and forming N classes of fragments. The first class has the lower limit zero and the last class has no upper limit. The intermediate class limits were obtained by averaging two successive flows of the sorted series of annual flows. The corresponding fragments were assigned to each class. The generated annual flows are then checked one-by-one for the class to which they belong, and disaggregated using the corresponding fragment. By selecting an appropriate fragment for each synthetic streamflow, the preservation of the historical monthly statistics improved remarkably. Santos (1983, p. 20) presented an explanation for such improvement: when there is a relationship between the within-year distribution of monthly flows of a given year and the total streamflow of that year, it must be taken into account when selecting the fragments to disaggregate the synthetic annual streamflows.

Although the method of fragments is capable of generating synthetic series of monthly flows, while generally preserving the monthly and annual statistics of the sample and the additive property, it has been known not to preserve the monthly serial correlation between the first month of a year and the last month of the previous year (Srikanthan and McMahon 1980a, 2001)

Porter and Pink (1991) reported that the use of the method of fragments resulted in the generation of monthly flows with a noticeable repetitive and cyclical pattern. Those authors proposed a method of synthetic fragments where the fragments were obtained from a generated monthly flow sequence. However, this method did not overcome the problem of the non-preservation of the serial correlation between the first month of a given year and the last month of the previous one (Srikanthan and McMahon 2010). Maheepala and Perera (1996) proposed a modification to the method of synthetic fragments and compared it to both the Porter and Pink (1991) and the Srikanthan and McMahon (1980a, 1982) variations of the method. The results showed that the modification preserved annual and monthly parameters satisfactorily, as well as monthly correlations across consecutive years.

Regarding the methods of synthetic fragments, Srikanthan and McMahon (2001) commented that Porter and Pink (1991) and Maheepala and Perera (1996) are not clear as to how the synthetic monthly data were generated and that, furthermore, it is not clear how they handle the problem of generating the right amount of zero monthly flows, which is one of the major problems in applying data generation models to ephemeral streams, according to Srikanthan and McMahon (1980a). This is an important aspect of the streamflow regime in mainland Portugal, particularly in the southern region.

Santos (1983) suggests that the fragments should be grouped into classes that may have more than one fragment. When disaggregating a given generated annual streamflow, a fragment is randomly selected from the class to which that annual flow belongs. Arsénio (2003) adopted this technique and suggests that, when disaggregating a synthetic annual streamflow series, the random selection of fragments from a given class is carried out without replacement, until the class becomes empty. At that point, all the fragments of that class are reintroduced into it. In order to study the effect that the number and amplitude of the classes of fragments have on the performance of the method, Arsénio (2003) used the method of fragments for disaggregating monthly flows considering different criteria regarding the number and size of the classes. Arsénio (2003) concluded that using classes of different sizes yielded better results, and that using a small number of classes (each one comprising a large number of fragments) produced the worst results. Arsénio (2003) also suggested that the definition of the classes of fragments be made by trial and error until an assembly of fragments is found that leads to a more effective preservation of the statistics of the samples. This approach constitutes a source of ambiguity to the application of this method. Recently, Silva (2010) and Silva and Portela (2010, 2011) applied the method of fragments to generate a large number of synthetic monthly flow series for 53 gauging stations in Portugal and obtained good results. Those authors defined the classes of fragments as probability intervals and thus avoided the trial-and-error approach proposed by Arsénio (2003). The objective of this study is to evaluate the performance of the method of fragments using different criteria for the definition of classes and the selection of fragments.

Following the brief theory review of the method of fragments presented in this Introduction, a methodology is proposed for disaggregating annual flows into monthly flows using the method of fragments, in which three criteria for the selection of fragments have been used. Next, the streamflow data are presented and an attempt is made to characterize relationship between the fragments and the historical annual flows. The methodology is applied and an overall evaluation and discussion of the obtained results is made. Finally, some relevant conclusions are drawn.

2 METHODOLOGY

2.1 General considerations

The methodology used to generate synthetic monthly streamflow series consists of a two-level approach, comprising a generation model for the annual synthetic series and, at the monthly level, three approaches of the method of fragments that differ in their definition of the classes and selection of fragments. The synthetic series were generated with a length equal to that of the corresponding sample.

The generation of synthetic streamflow series is a widely-used technique in the framework of studies on reservoir storage-yield design (McMahon and Adeloye 2005, p. 123). Using streamflow data from four Portuguese rivers, Guimarães (2005) and Guimarães and Santos (2011) studied the number of synthetic series, M, to generate from a given streamflow sample, in the context of water-supply reservoir design studies, and concluded that it should be 1200. This number was adopted in the current analyses.

According to Quintela (1967), the annual streamflow series of Portuguese rivers is an independent time series if the hydrological year (starting on 1 October) is adopted as the time step, provided that an analysis of the correlation structure of the series does not invalidate this assumption. Hence, the modelling of the annual streamflow series may use a non-deterministic probabilistic model.

2.2 Generation of annual streamflows

The model for generating annual flows consists of a random sampling of the log-Pearson Type III distribution. Because that distribution can only take values between 0 and $+\infty$, the generation of negative flows is naturally avoided. Furthermore, the log-Pearson III law has a flexible distribution, assuming a number of different shapes depending on the mean, variance and skewness of the sample, and, hence, the treatment of skewed data is assured (Chow *et al.* 1988, p. 375).

2.3 Monthly disaggregation of streamflows using the method of fragments

As mentioned in the Introduction, the method of fragments uses a purely deterministic approach to disaggregate synthetic annual flows into synthetic monthly flows. In this method, for a given year, k, the observed monthly flows are divided by the corresponding annual flow volume, X_k , and the resulting set of 12 standardized monthly flows constitutes the fragment pertaining to the year k, ϕ_k :

$$\phi_k = \frac{X_{k,j}}{X_k} = \begin{bmatrix} \frac{X_{k,1}}{X_k} & \frac{X_{k,2}}{X_k} & \cdots & \frac{X_{k,11}}{X_k} & \frac{X_{k,12}}{X_k} \end{bmatrix}$$
(1)

where $X_{k,j}$ are the monthly flows (j = 1, 2, ..., 12).

The application of the method of fragments requires that the fragments are formed and assembled into classes beforehand. For the constitution of the array of fragments, $[\phi]$, the annual flows must previously be ranked from smallest to largest:

$$[\phi] = \begin{bmatrix} \phi_1 \\ \vdots \\ \phi_k \\ \vdots \\ \phi_N \end{bmatrix} = \begin{bmatrix} \frac{X_{1,1}^*}{X_1^*} & \cdots & \frac{X_{1,12}^*}{X_1^*} \\ \vdots & \ddots & \vdots \\ \frac{X_{k,1}^*}{X_k^*} & \cdots & \frac{X_{k,12}^*}{X_k^*} \\ \vdots & \ddots & \vdots \\ \frac{X_{N,1}^*}{X_N^*} & \cdots & \frac{X_{N,12}^*}{X_N^*} \end{bmatrix}$$
(2)

where X_k^* represents the *k*th annual flow such that $X_{k-1}^* \leq X_k^* \leq X_{k+1}^*$, and X_{k+1}^* represents the monthly flows of that year.

After the array of fragments is formed, one has to select a fragment, ϕ_i , to disaggregate each synthetic annual flow, \hat{X}_i , into 12 monthly flows, $\hat{X}_{i,j}$, as follows:

$$\hat{X}_{i,j} = \phi_i \hat{X}_i \tag{3}$$

where i = 1, 2, ..., N and j = 1, 2, ..., 12.

According to Santos (1983, pp. 19–20), the way in which the fragments are selected for each synthetic annual flow is the most important feature of the method. In the state of the art involving the method of fragments, reviewed in the introduction, a number of different procedures are available for carrying out that selection, and this generally involves forming classes of fragments. In this paper, the method of fragments was implemented using three different criteria to define the classes of fragments and the selection of fragments. A description of each criterion follows (if pertinent, such description assumes that the annual flows are already ranked).

- Criterion 1 This criterion, which constitutes the initial concept of the method of fragments by Svanidze (1964) (Srikanthan and McMahon 1980a, 1980b), considers a single class ranging from 0 to +∞, the selection of fragments being random without replacement. The fragments are not replaced as they are selected in order to avoid a repetitive pattern in the generated monthly flow sequences. This procedure considers that the within-year distribution of streamflows is purely random, regardless of annual flows.
- **Criterion 2** The second criterion adopts the modification proposed by Srikanthan and McMahon (1980a, 1980b, 1982), in which they introduced a relationship between the fragments and the annual streamflow. This procedure considers one class per fragment, the lower limit of the first class being zero, the last class having no upper limit, and the remaining class limits being obtained by averaging each two consecutive values of the ranked series of annual streamflows. For each year of the synthetic series, the generated annual flow is checked to determine to which class it belongs. The selected fragment is the one associated with that class.
- Criterion 3 This procedure, used by Silva (2010) and Silva and Portela (2010, 2011), defines the classes of fragments as probability intervals with uniform amplitude of 10%. It considers classes with one or more fragments and proposes a robust and well-defined methodology for the definition of these classes, thus avoiding the ambiguous and time-consuming trial-and-error approach used by Arsénio (2003). Though additional research is still required regarding the effect of the amplitude of

the intervals in the performance of the criterion, the procedure implemented is as follows.

First, a series of nine equally-spaced nonexceedence probabilities of the annual streamflow, with an increment of 10%, is considered (10%, 20%, ..., 90%). Next, one estimates the annual flows, \breve{X}_m , with the aforementioned non-exceedence probabilities, by inverting the probability distribution function of the log-Pearson III law. This results in 10 equally probable classes with limits represented by $[0; \breve{X}_1[, [\breve{X}_1; \breve{X}_2[, [\breve{X}_3; \breve{X}_3[, ..., [\breve{X}_8; \breve{X}_9[, [\breve{X}_9; +\infty[.$

The fragments are distributed among the successive classes, while simultaneously verifying if there are any empty classes. Any empty classes that occur must be redefined. If the first class is empty, it is included in the next class; if the last class is empty, it is included in the previous class; if an intermediate class is empty, half of the probability step that defines it is attributed to each of the two bordering classes.

The fragments are randomly selected, without replacement. If any class runs out of fragments, it is refilled with the respective fragments.

It should be stressed that, while criteria 2 and 3 consider that there is a relationship between the annual streamflow and the within-year distribution of monthly flows, Criterion 1 considers that both features are independent.

Following the generation of the synthetic annual and monthly streamflow series, the results of the application of the method of fragments in accordance with each criterion were compared in terms of its ability to preserve the main seasonal statistical parameters of the historical samples, namely the mean, standard deviation, skewness coefficient and lag-one serial correlation of the monthly streamflow series.

3 STREAMFLOW DATA

The method of fragments, with the class definition and the fragment selection given by each criterion, was applied to obtain synthetic monthly flow series based on the streamflow samples from 34 gauging stations in unregulated Portuguese rivers. The names of the gauging stations and the main characteristics of the samples are presented in Table 1. The stations were selected from the set of 53 used by Silva (2010) and by Silva and Portela (2010, 2011), according to the following two rules: longest samples; even geographical distribution. By applying the above rules, 31 stations with at least 30 years of records were selected. This data set was complemented with three

| Sample no. | Gauging station (River) | Period of record (years) | A (km ²) | \overline{X} (10 ³ m ³) | \overline{H} (mm) | $s_X (10^3 \text{ m}^3)$ | <i>g</i> ^{<i>X</i>} (-) | <i>C_V</i> (-) |
|------------------|---|---|----------------------|--|---------------------|--------------------------|----------------------------------|--------------------------|
| 1 | Açude Saimilo (Dão) | 1939/40–1974/75 (36) | 1371 | 86884 | 634 | 400606 | 0.6426 | 0.4611 |
| 2 | Amieira (Degebe) | 1944/45–1990/91 (47) | 1474 | 214233 | 145 | 209450 | 0.9949 | 0.9777 |
| 3 | Bodega (Rib. Alportel) | 1952/53-1988/89 (37) | 132 | 30852 | 234 | 22401 | 0.6167 | 0.7261 |
| 4 | Cabriz (Rib. S. Paio) | 1966/67–1996/97 (31) | 17 | 13155 | 774 | 5738 | 0.6769 | 0.4362 |
| 5 | Caldas S. Gemil (Dão) | 1952/53-1989/90 (38) | 617 | 293547 | 476 | 166582 | 0.6629 | 0.5675 |
| 9 | Castanheiro (Tua) | 1958/89-2003/04 (46) | 3718 | 1401978 | 377 | 895045 | 1.0687 | 0.6384 |
| 7 | Castelo Bom (Côa) | 1957/58-2003/04 (47) | 897 | 311726 | 348 | 187310 | 0.4558 | 0.6009 |
| 8 | Castro Daire (Paiva) | 1945/46-2003/04 (59) | 291 | 214396 | 737 | 114522 | 0.8910 | 0.5342 |
| 6 | Cidadelhe (Côa) | 1955/56-2003/04 (49) | 1685 | 506950 | 301 | 336010 | 0.3863 | 0.6628 |
| 10 | Couto de Andreiros | 1974/75-1992/93 (19) | 244 | 46749 | 192 | 44017 | 0.8019 | 0.9415 |
| | (Rib. de Seda) | | | | | | | |
| 11 | Cunhas (Beça) | 1949/50-2005/06 (57) | 338 | 278348 | 824 | 140064 | 0.7109 | 0.5032 |
| 12 | Ermida-Corgo (Corgo) | 1956/57-2005/06 (50) | 291 | 258443 | 888 | 126489 | 0.8231 | 0.4894 |
| 13 | Flor da Rosa (Xamarra) | 1934/35-1965/66 (32) | 278 | 73727 | 265 | 54052 | 0.2039 | 0.7331 |
| 14 | Fragas da Torre (Paiva) | 1945/46-2005/06 (61) | 660 | 656081 | 994 | 331357 | 0.9442 | 0.5051 |
| 15 | Manteigas (Zêzere) | 1948/49-1995/96 (48) | 28 | 57638 | 2058 | 20145 | 0.8992 | 0.3495 |
| 16 | Monforte (Rib. Aviz) | 1955/56-1988/89 (34) | 136 | 31387 | 231 | 23590 | 0.6450 | 0.7516 |
| 17 | Mte. da Ponte (Cobres) | 1959/60 - 1993/94 (35) | 701 | 89087 | 127 | 85245 | 1.0678 | 0.9569 |
| 18 | Mte. dos Fortes (Rib. | 1961/62-1992/93 (32) | 288 | 74632 | 259 | 52283 | 0.6922 | 0.7005 |
| | Odeleite) | | | | | | | |
| 19 | Murça (Tinhela) | 1970/71-2003/04 (34) | 265 | 106080 | 400 | 78831 | 1.7679 | 0.7431 |
| 20 | Odivelas (Rib. Odivelas) | 1934/35-1969/70 (36) | 431 | 78752 | 183 | 57575 | 0.3522 | 0.7311 |
| 21 | Pte. Canavezes | 1955/56-1986/87 (32) | 3180 | 2195048 | 069 | 1057224 | 0.7741 | 0.4816 |
| | (Tâmega) | | | | | | | |
| 22 | Pte. Cavez (Tâmega) | 1957/58-2005/06 (49) | 1951 | 1012890 | 519 | 633543 | 1.1977 | 0.6255 |
| 23 | Pte. Freiria (Maior) | 1976/77–1989/90 (14) | 184 | 63111 | 343 | 43887 | 0.5154 | 0.6954 |
| 24 | Pte. Juncais (Mondego) | 1918/19-1990/91 (73) | 604 | 285150 | 472 | 172424 | 1.2867 | 0.6047 |
| 25 | Pte. Mucela (Alva) | 1938/39–1989/90 (52) | 666 | 404554 | 607 | 208138 | 0.3598 | 0.5145 |
| 26 | Pte. Nova (Almonda) | 1976/77 - 1989/90 (14) | 102 | 116204 | 1139 | 53017 | -0.4894 | 0.4562 |
| 27 | Pte. Sta. Clara-Dão | 1921/22-1988/89 (68) | 177 | 75824 | 428 | 51469 | 1.6676 | 0.6788 |
| | (Dão) | | | | | | | |
| 28 | Pte. Tábua (Mondego) | 1937/38–1978/79 (42) | 1552 | 654987 | 422 | 339001 | 0.5434 | 0.5176 |
| 29 | Pte. Vale Maior (Caima) | 1935/36-1972/73 (38) | 188 | 117572 | 625 | 70235 | 1.2203 | 0.5974 |
| 30 | Qta. das Laranjeiras | 1942/43-2005/06 (64) | 3464 | 851107 | 246 | 640723 | 1.0873 | 0.7528 |
| | (Sabor) | | | | | | | |
| 31 | Rebordelo (Rabaçal) | 1955/56-2002/03 (48) | 857 | 512542 | 598 | 302796 | 1.0300 | 0.5908 |
| 32 | Sta. Marta do Alvão | 1955/56-2005/06 (51) | 52 | 47996 | 923 | 21222 | 0.8840 | 0.4422 |
| | (Louredo) | | | | | | | |
| 33 | Vale Giestoso (Beça) | 1957/58-2005/06 (49) | 77 | 54215 | 704 | 33159 | 1.1520 | 0.6116 |
| 34 | Vinhais (Tuela) | 1955/56–1996/97 (42) | 455 | 360177 | 792 | 157043 | 0.5682 | 0.4360 |
| Note: A, watersh | ned area; \overline{X} , mean; \overline{H} , mean de | spth; s_X , standard deviation; g_X | , skewness coe | ficient; C_V , coeffi | cient of variatic | on of the annual flo | SWS. | |

Disaggregation modelling of monthly streamflows using a new approach of the method of fragments

additional stations (numbers 10, 23 and 26) that have considerably shorter samples, but that cover areas where there is little or no information. The location of the 34 stations is presented in Fig. 1. As is shown in Table 1, the mean annual flow depth of the samples ranges from 127 to 2058 mm, thus representing a wide variety of hydrological regimes.

The correlation structure of the historical samples of annual streamflow was analysed by computing the lag-one autocorrelation coefficients of the annual flows. It was found that, at the analysed gauging stations, the annual streamflow values constitute an independent time series, provided that the hydrological year (starting on 1 October) is adopted.

A visual analysis was made for each sample to determine whether there was a relationship between the annual flow and the within-year distribution of monthly flows. An example of such analysis is given in Fig. 2 for sample no. 2 (Amieira). In Fig. 2, for each year of the historical annual series (47 years for this sample), there is a set of 12 markers, displayed horizontally, representing the values that constitute the fragment for the annual streamflow indicated in the vertical axis. The colour of each marker indicates the ratio, R, between that monthly flow and the corresponding annual flow. If there was a relationship between the annual flow and the within-year distribution of monthly flows, one would expect the colours



Fig. 1 Map of Portugal showing location of the 34 gauging stations.



Fig. 2 Visual analysis of the relationship between the fragments and the annual streamflows of sample no. 2 (Amieira).

of the markers to vary at the same rate at which the annual flow increases or decreases, which does not happen: the marker colours for the different months do not exhibit discernible patterns.

For the set of 34 gauging stations, the visual analvses based on diagrams such as that shown in Fig. 2 were not conclusive as to whether or not there is any relationship between the annual flow and the withinyear distribution of monthly flows. Furthermore, it is an ambiguous analysis, since it relies on subjective interpretation. Santos (1983, p. 20) presents a more objective approach for evaluating whether or not there is a relationship between the within-year distribution of monthly flows and the annual flows of the streamflow sample. The approach is exemplified in Fig. 3. For a given sample, it consists of dividing the monthly flows of each year, $X_{i,j}$, by the mean monthly flow of that year, $\overline{X_i}$. Considering that the total time is one, i.e. each month is 1/12 in annual time units, the within-year irregularity factor, d, is the area above $X_{i,i}/\overline{X_i} = 1$. The area below the line is a natural flow regulation factor, ψ . Naturally, $\psi + d = 1$.

For each streamflow sample, the d values were plotted against the annual flows and the correlation



Fig. 3 Approach for finding the within-year irregularity factor, *d*.

coefficient, r, between the two variables is determined. Figure 4 shows the analysis for sample no. 2 (on the horizontal axis, the annual flows were divided by the mean annual flow of the sample). Table 2 shows the correlation coefficient, r, obtained for each sample.

The procedure proposed by Yevjevich (1972, pp. 237–242) was used to determine whether a given value of *r* differs significantly from zero. A significance level of $\alpha = 5\%$ was considered. It was found that in 11 of the samples, the correlation coefficients are significantly different from zero (indicated in bold in Table 2).



Fig. 4 Relationship between the d values and the annual flows in sample no. 2 (Amieira).

Table 2 Correlation coefficients, r, between the d values and the historical annual flows.

| Sample no. | Gauging station (River) | r* | | |
|------------|-----------------------------------|--------|--|--|
| 1 | Açude Saimilo (Dão) | 0.163 | | |
| 2 | Amieira (Degebe) | 0.256 | | |
| 3 | Bodega (Rib. Alportel) | -0.148 | | |
| 4 | Cabriz (Rib. S. Paio) | -0.042 | | |
| 5 | Caldas S. Gemil (Dão) | 0.189 | | |
| 6 | Castanheiro (Tua) | 0.418 | | |
| 7 | Castelo Bom (Côa) | 0.094 | | |
| 8 | Castro Daire (Paiva) | 0.365 | | |
| 9 | Cidadelhe (Côa) | -0.064 | | |
| 10 | Couto de Andreiros (Rib. de Seda) | -0.192 | | |
| 11 | Cunhas (Beça) | 0.241 | | |
| 12 | Ermida-Corgo (Corgo) | 0.218 | | |
| 13 | Flor da Rosa (Xamarra) | -0.134 | | |
| 14 | Fragas da Torre (Paiva) | 0.415 | | |
| 15 | Manteigas (Zêzere) | 0.027 | | |
| 16 | Monforte (Rib. Aviz) | -0.127 | | |
| 17 | Mte. da Ponte (Cobres) | -0.116 | | |
| 18 | Mte. dos Fortes (Rib. Odeleite) | -0.118 | | |
| 19 | Murça (Tinhela) | 0.542 | | |
| 20 | Odivelas (Rib. Odivelas) | -0.029 | | |
| 21 | Pte. Canavezes (Tâmega) | 0.253 | | |
| 22 | Pte. Cavez (Tâmega) | 0.438 | | |
| 23 | Pte. Freiria (Maior) | 0.529 | | |
| 24 | Pte. Juncais (Mondego) | 0.233 | | |
| 25 | Pte. Mucela (Alva) | 0.097 | | |
| 26 | Pte. Nova (Almonda) | 0.701 | | |
| 27 | Pte. Sta. Clara-Dão (Dão) | 0.519 | | |
| 28 | Pte. Tábua (Mondego) | 0.061 | | |
| 29 | Pte. Vale Maior (Caima) | -0.489 | | |
| 30 | Qta. das Laranjeiras (Sabor) | 0.399 | | |
| 31 | Rebordelo (Rabaçal) | 0.276 | | |
| 32 | Sta. Marta do Alvão (Louredo) | 0.170 | | |
| 33 | Vale Giestoso (Beça) | 0.554 | | |
| 34 | Vinhais (Tuela) | 0.323 | | |

* The correlations in **bold** are significantly different from zero (for $\alpha = 5\%$).

Notwithstanding the prior analyses, the authors believe that the relationship between annual streamflow and the within-year distribution of monthly streamflow is not particularly strong in Portuguese rivers. This conclusion seems to be in accordance with the results presented in Table 2, as even for the cases where the correlation differs significantly from zero, the correlation coefficients are relatively small.

4 RESULTS AND DISCUSSION

The quality of the method of fragments, based on the three criteria presented in Section 2.3, was assessed by comparing monthly statistical parameters, namely the mean, standard deviation, skewness coefficient and lag-one serial correlation coefficient. Each statistical parameter evaluated based on a given sample was compared to the average of 1200 values of that parameter obtained from an equal number of synthetic flow series, each of equal length to that of the corresponding sample, as previously mentioned. Three sets of 1200 synthetic flow series were considered: one for each of the three different criteria used in implementing the method of fragments.

The comparison was made by plotting the relative error (%) of the mean and standard deviation (SD) estimates, and the absolute error (-) of the skewness and serial correlation estimates, against the months. It should be noted that this analysis takes into account the hydrological year in Portugal, and so it starts in October.

The absolute and relative errors were calculated as follows, where Gen denotes the mean of the 1200 statistics estimated from the generated synthetic series and Hist the same statistic estimated from the corresponding historical sample:

Absolute error = Gen - Hist (4)

Relative error =
$$\frac{\text{Gen} - \text{Hist}}{\text{Hist}} \times 100$$
 (5)

Figure 5 shows an example of the results obtained for sample no. 11 (Cunhas). Since it is not feasible to represent the analogous figures for the remaining samples here, the sums of the absolute values of the relative and absolute errors were computed and are shown in Table 3.

The results of Fig. 5 and Table 3 show that, although the analyses carried out, as exemplified in Figs 2 and 4, were not conclusive as to whether there is a relationship between the annual streamflow and the within-year distribution of monthly flows, the consideration of such a relationship when selecting the fragments greatly improves the performance of the method of fragments; that is, criteria 2 and 3 perform much better than Criterion 1. They also show that, generally, Criterion 3 performs better than Criterion 2. In fact, as highlighted in Table 3, from the 34 samples and depending on the monthly parameter, the best results (smallest sum of errors) were obtained for Criterion 1 in zero (SD and lag-1 serial corr.) to three (mean) samples, for Criterion 2 in three (SD) to eight (lag-1 serial corr.) samples, and for Criterion 3 in 26 (lag-1 serial corr.) to 31 (SD) samples. Furthermore, when Criterion 3 does not show the best performance,



Fig. 5 Relative and absolute errors in the monthly parameters of sample no. 11 (Cunhas) (Hist. denotes the results obtained from the historic sample and Criterion 1 to 3, those from the method of fragments based on the different criteria).

| Sample no. | Mean (%) | | | SD (%) | | | Skew. coeff. (-) | | | Lag-1 serial corr. (-) | | |
|------------|----------|--------|--------|---------|--------|--------------|------------------|-------|--------|------------------------|-------|--------|
| | Crit 1 | Crit2 | Crit 3 | Crit 1 | Crit2 | Crit 3 | Crit 1 | Crit2 | Crit 3 | Crit 1 | Crit2 | Crit 3 |
| 1 | 50.42 | 52.08 | 22.71 | 180.39 | 103.30 | 66.25 | 5.48 | 2.92 | 2.20 | 0.73 | 0.82 | 0.51 |
| 2 | 1065.18 | 366.79 | 131.06 | 2600.94 | 390.60 | 211.57 | 17.49 | 9.49 | 6.22 | 3.34 | 2.88 | 1.17 |
| 3 | 448.47 | 157.00 | 119.38 | 807.77 | 143.57 | 84.42 | 6.95 | 7.66 | 3.30 | 1.38 | 0.82 | 0.49 |
| 4 | 71.05 | 70.11 | 43.19 | 226.06 | 125.43 | 33.93 | 6.42 | 2.13 | 1.41 | 0.91 | 0.94 | 0.63 |
| 5 | 73.27 | 54.69 | 33.21 | 266.48 | 82.05 | 50.22 | 6.08 | 3.38 | 1.87 | 1.40 | 0.81 | 0.58 |
| 6 | 242.95 | 110.49 | 34.52 | 726.90 | 138.13 | 72.20 | 6.30 | 3.88 | 2.55 | 1.34 | 0.97 | 0.62 |
| 7 | 197.02 | 89.65 | 65.57 | 417.40 | 125.68 | 69.64 | 7.94 | 4.97 | 3.37 | 1.18 | 0.98 | 0.51 |
| 8 | 145.51 | 56.65 | 39.89 | 468.68 | 82.66 | 77.35 | 7.57 | 3.61 | 1.78 | 1.73 | 0.90 | 0.80 |
| 9 | 377.90 | 152.09 | 132.12 | 917.17 | 173.33 | 80.61 | 8.88 | 3.66 | 2.77 | 1.49 | 0.72 | 0.84 |
| 10 | 764.05 | 315.23 | 290.24 | 1752.13 | 473.58 | 428.77 | 8.38 | 4.50 | 4.59 | 2.00 | 1.72 | 1.17 |
| 11 | 151.29 | 46.71 | 14.57 | 429.17 | 102.71 | 30.49 | 6.70 | 3.96 | 2.44 | 1.50 | 1.02 | 0.64 |
| 12 | 121.47 | 69.64 | 47.10 | 353.62 | 149.43 | 68.65 | 4.94 | 4.54 | 2.53 | 1.76 | 1.26 | 0.90 |
| 13 | 3988.00 | 386.17 | 452.41 | 5795.19 | 291.70 | 369.21 | 11.62 | 8.08 | 6.73 | 1.32 | 3.16 | 1.16 |
| 14 | 121.94 | 55.97 | 26.17 | 419.77 | 77.20 | 47.25 | 6.77 | 2.89 | 2.15 | 1.86 | 0.91 | 0.85 |
| 15 | 45.93 | 72.31 | 55.37 | 137.22 | 98.82 | 46.01 | 4.00 | 2.44 | 1.67 | 1.55 | 1.29 | 0.96 |
| 16 | 418.73 | 204.60 | 131.14 | 836.22 | 226.06 | 149.61 | 10.31 | 5.31 | 3.91 | 2.17 | 1.95 | 1.47 |
| 17 | 5768.00 | 235.04 | 120.74 | 7380.74 | 279.22 | 101.57 | 10.40 | 3.82 | 2.81 | 1.71 | 0.79 | 0.69 |
| 18 | 456.91 | 100.82 | 202.26 | 789.56 | 202.75 | 196.77 | 7.27 | 4.28 | 6.11 | 1.86 | 0.64 | 0.64 |
| 19 | 251.49 | 64.82 | 79.23 | 834.88 | 105.90 | 106.57 | 8.74 | 3.19 | 3.80 | 1.54 | 0.87 | 0.95 |
| 20 | 1661.44 | 308.88 | 267.83 | 1949.25 | 358.47 | 293.96 | 10.27 | 4.17 | 4.69 | 1.15 | 1.16 | 0.94 |
| 21 | 128.20 | 63.04 | 51.22 | 481.65 | 132.12 | 66.87 | 6.10 | 4.03 | 1.87 | 1.45 | 1.08 | 0.62 |
| 22 | 200.62 | 79.61 | 50.23 | 630.12 | 120.59 | 83.20 | 6.91 | 4.53 | 3.75 | 1.70 | 0.81 | 0.86 |
| 23 | 284.65 | 126.28 | 73.09 | 679.24 | 219.90 | 158.71 | 5.45 | 5.09 | 4.62 | 1.34 | 0.96 | 0.63 |
| 24 | 114.77 | 58.87 | 18.28 | 239.86 | 75.45 | 43.34 | 7.79 | 5.53 | 2.97 | 1.00 | 0.78 | 0.47 |
| 25 | 159.44 | 88.19 | 98.11 | 711.97 | 76.34 | 137.21 | 11.20 | 2.49 | 2.74 | 1.62 | 0.94 | 0.96 |
| 26 | 180.43 | 178.61 | 120.49 | 608.14 | 181.94 | 178.80 | 6.27 | 4.51 | 3.36 | 1.99 | 2.56 | 1.71 |
| 27 | 159.75 | 75.48 | 47.71 | 381.72 | 108.28 | 96.03 | 10.55 | 5.41 | 4.92 | 1.90 | 1.05 | 0.81 |
| 28 | 73.27 | 86.27 | 72.79 | 198.59 | 127.14 | 75.04 | 4.84 | 3.10 | 2.00 | 1.38 | 1.27 | 1.11 |
| 29 | 85.13 | 186.51 | 155.95 | 236.00 | 110.06 | 98.98 | 9.21 | 6.15 | 3.00 | 1.87 | 0.66 | 0.81 |
| 30 | 263.80 | 140.35 | 67.37 | 578.77 | 196.82 | 121.32 | 7.74 | 4.38 | 2.53 | 0.85 | 0.64 | 0.75 |
| 31 | 210.43 | 53.54 | 53.78 | 637.03 | 117.12 | 71.69 | 7.94 | 4.31 | 2.76 | 1.81 | 0.93 | 0.91 |
| 32 | 106.62 | 57.49 | 34.87 | 291.17 | 106.24 | 44.10 | 6.79 | 3.26 | 2.89 | 1.60 | 0.74 | 0.72 |
| 33 | 206.98 | 61.28 | 54.95 | 888.95 | 90.73 | 77.00 | 9.14 | 3.45 | 2.85 | 2.05 | 0.84 | 0.89 |
| 34 | 125.92 | 78.01 | 75.78 | 335.28 | 173.72 | 104.95 | 3.85 | 6.40 | 4.15 | 1.50 | 0.96 | 0.73 |

Table 3 Sum of the absolute values of the errors of the monthly parameters–relative errors for the means and standard deviations (SD), and absolute errors for the skewness coefficients and lag-1 serial correlations^{*}.

* The values in **bold** identify the criterion with the better performance.

it is never the least well-performing, presenting error measures very close to those of the best criterion.

In order to further analyse the difference in performance of criteria 2 and 3, box plots were drawn of the generated values of the analysed monthly parameters and these were compared with the historical parameters. In those plots, the whiskers have a maximum length of $1.5 \times IQR$ (interquartile range) and the values outside the whiskers are considered to be outliers (Robbins 2005, p. 91). Figures 6–8 show the comparison between the performance of the two methods for samples 2, 9 and 30. The figures pertaining to the remaining samples, though omitted here, showed similar results: in general, Criterion 3 slightly, but consistently, outperforms Criterion 2.

The latter results show that the consideration of fewer, wider classes (with more than one fragment),

and the introduction of randomness when selecting a fragment from a class—as implemented by Criterion 3—improves the performance of the method of fragments, as generally denoted by the smaller amplitudes of the boxes and whiskers. The higher variability of the results produced by Criterion 2, though not yet totally understood, may be related to the fact that using N classes reduces the probability of applying some of the fragments, especially those with very narrow classes, resulting from annual flows that are very close together.

The results also confirm a well-known limitation of the method of fragments: the non-preservation of the serial correlation between the first month of one year and the last month of the previous year (Srikanthan and McMahon 1980a, 2001). In fact, Figs 6–8 show that, for October, this coefficient is



Fig. 6 Comparison of the monthly parameters of the monthly historical and generated synthetic series of sample no. 2 (Amieira) (results obtained by using criteria 2 and 3 in the method of fragments).



Fig. 7 Comparison of the monthly parameters of the monthly historical and generated synthetic series of sample no. 9 (Cidadelhe) (results obtained by using criteria 2 and 3 in the method of fragments).



Fig. 8 Comparison of the monthly parameters of the monthly historical and generated synthetic series of sample no. 30 (Qta das Laranjeiras) (results obtained by using criteria 2 and 3 in the method of fragments).

consistently close to zero in the synthetic series. We believe that the method of fragments (under the formulations used herein) can only preserve the first lag-one serial correlation if the historical one is close to zero itself. This can be verified for sample no. 9 (Cidadelhe) in Fig. 7.

5 CONCLUSIONS

The method of fragments was used to disaggregate synthetic annual streamflow values into synthetic monthly streamflow series using three different criteria to define the classes of fragments and to select the fragments:

- a single class with fragments selected randomly and without replacement; accordingly all the fragments are used (Svanidze 1964);
- (2) as many classes as the number of fragments, with automatic replacement of each used fragment (Srikanthan and McMahon 1980b, 1980a, 1982); and
- (3) classes defined based on the assumption of equally probable width; the fragments are selected randomly and the replacement only occurs when the class to which they belong becomes empty; not necessarily all the fragments are used (Silva 2010, Silva and Portela 2010, 2011).

The different approaches were applied to 34 streamflow samples from gauging stations geographically spread throughout mainland Portugal. The performance of each approach was evaluated in terms of the capacity of the synthetic monthly series to preserve the main monthly statistical parameters of the historical samples. The results show that while criteria 2 and 3 perform much better than Criterion 1, Criterion 3 generally, outperforms Criterion 2. In fact, for most of the 34 samples (between 26 and 31 samples, depending on the monthly parameter), Criterion 3 ensures the smallest error measures. We stress that the results confirmed a well-known limitation of the method of fragments: the non-preservation of the serial correlation between the last month of one year and the first month of the following one.

Given the results described above, we recommend the method of fragments, integrated with he procedure developed by Silva (2010) and Silva and Portela (2010, 2011) for defining the classes and the selection of fragments (i.e. Criterion 3) for the disaggregation modelling of monthly streamflows in mainland Portugal.

The new approach of the method of fragments (Criterion 3) still provides opportunities for further development, such as: analysis of the effect of the amplitude of the probability intervals that define the classes of fragments in its performance; investigating the reason why it generally reduces the variability of the seasonal parameters of the generated series, compared to another approach (Criterion 2); the development of procedures for improving the preservation of the serial correlation between the last month of one year and the first month of the following one; and the possibility of its application to other time steps, namely daily.

REFERENCES

- Arsénio, J.J.G., 2003. *Gestão de Albufeiras: avaliação do potencial dos sistemas com recurso a séries sintéticas de escoamentos* (in Portuguese). Thesis (Master), Instituto Superior Técnico, Lisbon, Portugal.
- Chow, V.T., Maidment, D., and Mays, L., 1988. Applied hydrology. Singapore: McGraw-Hill International.
- Guimarães, R.C.P.D.C., 2005. Simulação no Dimensionamento e Gestão de Sistemas de Recursos Hídricos. Geração de Séries Sintéticas de Escoamento (in Portuguese). Thesis (PhD), Universidade de Évora, Évora, Portugal.
- Guimarães, R. and Santos, E.G., 2011. Principles of stochastic generation of hydrologic time series for reservoir planning and design: a case study. *Journal of Hydrologic Engineering*, 1, 237.
- Maheepala, S. and Perera, B., 1996. Monthly hydrologic data generation by disaggregation. *Journal of Hydrology*, 178 (1–4), 277–291.
- McMahon, T. and Adeloye, A., 2005. *Water resources yield*. Littleton, CO: Water Resources Publications.
- Ndiritu, J., 2011. A variable-length block bootstrap method for multisite synthetic streamflow generation. *Hydrological Sciences Journal*, 56 (3), 362–379.
- Porter, J.W. and Pink, B.J., 1991. Method of synthetic fragments for disaggregation in stochastic data generation. In: International Hydrology and Water Resources Symposium, 2–4 October 1991, Perth, Australia. Barton, ACT: Institute of Engineers, 781–786.
- Quintela, A., 1967. Recursos de águas superficiais em Portugal Continental (in Portuguese). Thesis (PhD), Instituto Superior Técnico, Lisbon, Portugal.
- Robbins, N.B., 2005. Creating more effective graphs. Hoboken, NJ: Wiley-Interscience.
- Salas, J.D., Delleur, J.W., Yevjevich, V., and Lane, W.L., 1980. Applied modeling of hydrologic time series. 2nd ed. Littleton, CO: Water Resources Publications.
- Santos, E.G., 1983. Disaggregation modeling of hydrologic time series. Thesis (PhD), Colorado State University, Colorado, USA.
- Santos, E.G. and Salas, J.D., 1992. Stepwise disaggregation scheme for synthetic hydrology. *Journal of Hydraulic Engineering*, 118 (5), 765–784.
- Silva, A.T.C. de F., 2010. Design of the storage capacity of artificial reservoirs. Thesis (Master), Instituto Superior Técnico, Lisbon, Portugal.

- Silva, A.T. and Portela, M.M., 2010. Geração de séries sintéticas de escoamentos anuais e mensais: aplicação extensiva a Portugal Continental (in Portuguese). *Revista Recursos Hídricos*, 31 (2), 27–41.
- Silva, A.T. and Portela, M.M., 2011. Generation of monthly synthetic streamflow series based on the method of fragments. In: Water Resources Management 2011—6th International Conference on Sustainable Water Resources Management, 23–24 May 2011, Riverside, CA. Hampshire: Wessex Institute of Technology.
- Srinkanthan, R. and McMahon, T.A., 1980a. Stochastic generation of monthly flows for ephemeral streams. *Journal of Hydrology*, 47 (1–2), 19–40.
- Srinkanthan, R. and McMahon, T.A., 1980b. Stochastic time series modelling of arid zone streamflows. *Hydrological Sciences Bulletin*, 25 (4), 423–434.

- Srinkanthan, R. and McMahon, T.A., 1982. Stochastic generation of monthly streamflows. *Journal of the Hydraulics Division of the American Society of Civil Engineers*, 108 (3), 419–441.
- Srinkanthan, R. and McMahon, T.A., 2001. Stochastic generation of annual, monthly and daily climate data: a review. *Hydrology* and Earth System Sciences, 5 (4), 653–670.
- Svanidze, G.G., 1964. Osnovy rascheta regulirovaniya rechnogo stoka metodom Monte-Karlo. *Izd. Akad. Nauk GSSR*, Tbilisi (in Russian).
- Svanidze, G.G., 1980. Mathematical modeling of hydrologic series; for hydroelectric and water resources computations. Littleton, CO: Water Resources Publications.
- Yevjevich, V., 1972. Probability and statistics in hydrology. Fort Collins, CO: Water Resources Publications.

14