Hydrol. Earth Syst. Sci. Discuss., 7, 5099–5130, 2010
www.hydrol-earth-syst-sci-discuss.net/7/5099/2010/
doi:10.5194/hessd-7-5099-2010

Hydrology and Earth System Sciences Discussions

Examination of homogeneity of selected Irish pooling groups

S. Das and C. Cunnane

Department of Engineering Hydrology, National University of Ireland, Galway, Ireland

Received: 1 July 2010 – Accepted: 9 July 2010 – Published: 30 July 2010

Correspondence to: S. Das (s.das1@nuigalway.ie)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

In regional flood frequency estimation, a homogeneous pooling group of sites leads to a reduction in the error of quantile estimators which is the main aim of a regional flood frequency analysis. Examination of the homogeneity of regions/pooling groups is usually based on a statistic that relates to the formulation of a frequency distribution model, e.g. the coefficient of variation (Wiltshire, 1986; Fill and Stedinger, 1995) and/or skew coefficient, their L-moment equivalents (Chowdhury et al., 1991; Hosking and Wallis, 1997) or of dimensionless quantiles such as the 10-yr event (Dalrymple, 1960; Lu and Stedinger, 1992). Hosking and Wallis (1993, 1997) proposed homogeneity tests based on L-moment ratios such as L-CV alone (H1) and L-CV & L-skewness jointly (H2) which were also recently investigated by Viglione et al. (2007). In this paper a study, based on annual maximum series obtained from 85 Irish gauging stations, examines how successful a common method of identifying pooling group membership is in selecting groups that actually are homogeneous. Each station has its own unique pooling group selected by use of a Euclidean distance measure in catchment descriptor space, commonly denoted d_{ij} and with a minimum of 500 station years of data in the pooling group, which satisfies the 5T rule (FEH, 1999, 3, p. 169) for the 100 yr quantile. It was found that d_{ij} could be effectively defined in terms of catchment area, mean rainfall and baseflow index. The sampling distribution of L-CV (t_2) in each pooling group and the 95% confidence limits about the pooled estimate of t_2 are obtained by simulation. The t_2 values of the selected group members are compared with these confidence limits both graphically and numerically. Of the 85 stations, only 1 station's pooling group members have all their t_2 values within the confidence limits, while 7, 33 and 44 of them have 1, 2 or 3 or more, t_2 values outside the confidence limits. The outcomes are also compared with the heterogeneity measures H1 and H2. The H1 values show an upward trend with the ranges of t_2 values in the pooling group whereas the H2 values do not show any such dependency. A selection of 27 pooling groups, found to be heterogeneous, were further examined with the help of box-plots

of catchment descriptor values and one particular case is considered in detail. Overall the results show that even with a carefully considered selection procedure, it is not certain that perfectly homogeneous pooling groups are identified. As a compromise it is recommended that a group containing more than 2 values of t_2 outside the confidence limits should not be considered homogeneous.

1 Introduction

It is widely accepted that a short annual flood (AM) series is inadequate for the estimation of design floods of large return periods. Regionalization (FSR, 1975), i.e. pooling analysis (FEH, 1999), is therefore used to provide a framework for design floods. In pooling analysis flood data are pooled from other gauging stations that possess similar hydrological behaviours to the at-site station. A very common way to implement regional/pooling is the index flood method proposed by Dalrymple (1960). Flood estimation based on this approach involves derivation of a growth curve which shows the relation between X_T and the return period T where $X_T = Q_T/Q_I$ and Q_I is the index flood at the site of interest. Generally the mean (FSR, 1975) or median (FEH, 1999) of the at-site AM flood series is taken as the index flood. It is assumed that the $X_T - T$ relation is the same at all sites in a homogeneous pooling group. The identification of a homogeneous pooling group is therefore important in pooling analysis. Lettenmaier et al. (1987); Stedinger and Lu (1995) and Hosking and Wallis (1997) among other researchers have demonstrated that a successful pooling analysis requires a homogeneity criterion to be satisfied.

An examination of homogeneity is normally used to assess whether a proposed group of sites is homogeneous or not. Examination of the homogeneity of regions/pooling groups is usually based on a statistic that relates to the formulation of a frequency distribution model, e.g. the coefficient of variation, CV (Wiltshire, 1986; Fill and Stedinger, 1995) and/or skew coefficient, g, their L-moment equivalents (Chowdhury et al., 1991; Hosking and Wallis, 1997) or of dimensionless quantiles such as the

HESSD
7, 5099–5130, 2010

10-yr event (Dalrymple, 1960; Lu and Stedinger, 1992). Hosking and Wallis (1993, 1997) proposed homogeneity tests based on L-moment ratios such as L-CV alone (H1) and L-CV & L-skewness jointly (H2) which are widely used in flood frequency analysis although the former one is recommended by these authors for having better power to discriminate between homogeneous and heterogeneous regions. Very recently, a similar conclusion has been drawn by Viglione et al. (2007) when they compared several homogeneity tests. They stated that H1 test is ahead of all others when the L-skewness is lower than 0.23. They further concluded that the H2 as a homogeneity test lacks power. These findings certainly indicate that the heterogeneity among the sites in a group is mainly due to variations in the sample L-CVs. However, one of the main assumptions of these tests is that the true regional distribution is kappa. For that reason and others Hosking and Wallis (1997) recommended that though the heterogeneity statistic is constructed like a significance test it should not be used in that way. They further stated that (Hosking and Wallis, 1997, p. 70)

> *. . . a significance test is of doubtful utility anyway, because even a moderately heterogeneous region can provide quantile estimates of sufficient accuracy for practical purposes. Thus a test of exact homogeneity is of little interest.*

In this paper a graphical way of examining the homogeneity of a pooling group is presented which is based on L-CV , i.e. t_2. The main idea behind the approach is the comparison of the variability of t_2 from each site in the pooling group with that expected (un-weighted average pooled t_2) supposing the differences between sites to be due to sampling error. The pooling groups are identified by the Region of Influence (ROI) approach. The population distribution for this purpose is selected based on the descriptive ability of the summary statistics of the region concerned.

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

7, 5099–5130, 2010

Homogeneity test

S. Das and C. Cunnane

2 Formation of pooling groups using Region Of Influence approach

The Region Of Influence (ROI) approach of formation of a pooling group is considered to be the most appropriate and meaningful way of delineating a pooling group. The technique developed by Burn (1990), involves the identification of a region of influence i.e. a separate pooling group for each gauging station in a region. The identification of a pooling group consists of selecting stations that are hydrologically similar to the site of interest. Similarity is measured generally by a Euclidean distance measure in catchment descriptor space. Careful consideration is necessary as to which form of catchment descriptors are to be used in a ROI method of pooling analysis.

2.1 Choice of catchment descriptors on effectiveness of ROI distance measures

The general form of the similarity measure used for selecting members of a pooling group is defined by

$$d_{ij} = \sqrt{\sum_{k=1}^{n} W_k \left(X_{k,i} - X_{k,j}\right)^2} \quad (1)$$

where d_{ij} is the weighted Euclidean distance from site j to site i; n is the number of attribute variables; $X_{k,i}$ is the value of the kth variable at the ith site and W_k is the weight applied to attribute k, reflecting its relative importance. The subscript i denotes to the subject site and the subscript j denotes to the jth pooled site.

In choosing a distance measure d_{ij} a decision has to be made about which catchment descriptors are to be included in the distance measure and what weightings are to be applied to them and whether logarithms or other transformations are to be used. The FEH (1999) provided a number of useful maxims for choosing a distance measure. It recommended not to use at-site flood statistics (e.g. CV, g) as pooling variables because this might well result in groups consisting of sites that have experienced similar floods in recent history. Neither could such site flood statistics be used for ungauged catchments. For selecting the final set of pooling variables, FEH used pooled uncertainty measure (PUM) which is a weighted average of the differences between each at-site growth factor and the pooled growth factor measured on a logarithmic scale. For Irish conditions, four catchment descriptors have been selected as the potential pooling variables from a pool of nineteen catchment descriptors made available by the Irish Office of Public Works (OPW). These are AREA (catchment area), SAAR (standard average annual rainfall) , BFI (baseflow index) and FARL (flow attenuation by reservoir and lake). These variables have been found to be effective in explaining the observed variation in Qmed values in the FSU (2009) regression studies. One such formulation is as follows

$$d_{ij} = \sqrt{\left(\frac{\ln A_i - \ln A_j}{\sigma_{\ln A}}\right)^2 + \left(\frac{\ln \mathrm{SAAR}_i - \ln \mathrm{SAAR}_j}{\sigma_{\ln \mathrm{SAAR}}}\right)^2 + \left(\frac{\mathrm{BFI}_i - \mathrm{BFI}_j}{\sigma_{\mathrm{BFI}}}\right)^2} \quad (2)$$

The objective is to find which combinations of these descriptors, which are listed in Table 2, lead to pooling groups which are most effective at exploiting the information about the flood distribution contained in the pooling groups. This is assessed by means of a Monte Carlo simulation procedure, which is described below. The GEV distribution is considered to be representative of what is appropriate in Irish conditions. Hosking and Wallis (1997, p. 93) suggested not to use the observed sample L-moment ratios as the population L-moment ratios of the simulated region because this would yield a simulated region that has much more heterogeneity than the actual data. Castellarin et al. (2001) addressed the issue by using a region of influence approach to estimate the at-site population values of t_2 and t_3. A similarity measure based on at-site flood statistics is used to form a group of sites for a subject site and its population values of t_2 and t_3 are considered as the corresponding pooled estimate of t_2 and t_3 for the group. Later, Gaal et al. (2008) adopted this approach in their study. A similar kind of

HESSD
7, 5099–5130, 2010

approach is used here with a similarity measure defined as

$$d_{ij} = \sqrt{\left(\frac{t_{2,i} - t_{2,j}}{\sigma_{t_2}}\right)^2 + \left(\frac{t_{3,i} - t_{3,j}}{\sigma_{t_3}}\right)^2} \quad (3)$$

which is independent of the descriptor variables being considered in Table 2. The estimated pooled values of t_2 and t_3 are then used as population values for each site in step 2 of the simulation procedure. The steps of the simulation procedure to selecting variables are described as follows.

1. The gauging stations in the subject site's pooling group are identified using the d_{ij} values of Eq. (1) for a set of catchment descriptors having a minimum of 5T station years of data in the pooling group.

2. Random samples are drawn from GEV populations for the subject site and for each site in the pooling group. For each site the sample size is taken as being equal to the length of the observed historical record at the site and the parameters are estimated from the site t_2 and t_3 values obtained using the procedure described above (e.g. using Eq. 3).

3. The t_2 and t_3 values are obtained for each sample in the pooling group and the weighted average of these is calculated to represent the pooled t_2 and t_3 values.

4. The pooled t_2 and t_3 values are then used to determine the pooling group's GEV growth curve parameters k and β using Eqs. (B3) and (B5).

5. The subject site's $\hat{X}_T$ value is calculated for T = 50 and 100 yr, respectively using Eq. (B2)

6. Steps 2 to 5 are repeated 10 000 times to provide 10 000 values of $\hat{X}_T$ and the

Discussion Paper | Discussion Paper | Discussion Paper | Discussion Paper |

HESSD

7, 5099–5130, 2010

Homogeneity test

S. Das and C. Cunnane

$RMSE_T$ and $BIAS_T$ are calculated for the subject site by the following equations:

$$\mathrm{RMSE}_T[\%] = \frac{1}{M}\sum_{i=1}^{M}\sqrt{\frac{1}{S}\sum_{s=1}^{S}\left(\frac{\hat{X}_{i,s}^{T}-X_{i}^{T}}{X_{i}^{T}}\right)^{2}}\cdot 100 \quad (4)$$

$$\mathrm{BIAS}_T[\%] = \frac{1}{M}\sum_{i=1}^{M}\frac{1}{S}\sum_{s=1}^{S}\left(\frac{\hat{X}_{i,s}^{T}-X_{i}^{T}}{X_{i}^{T}}\right)\cdot 100 \quad (5)$$

where $\hat{X}_{i,s}^{T}$ is the estimated T-year growth factor at a site i at the sth repetition; X_{i}^{T} is the assumed true T-year growth factor at site i; M is the number of sites in the pooling group and S is the number of repetitions.

$RMSE_T$ and $BIAS_T$ defined in the simulation procedure has been evaluated at 50 and 100-yr return periods for each site. The eight combinations is listed in Table 2 of the four variables have been tested based on $RMSE_T$ (primarily). In all, 85 stations have been considered for the study. The data sets that have been used in the study are summarized in Table 1. For each of these sites, a pooling group was selected from the 85 stations. Initially in the simulation procedure all weights W_k in Eq. (1) were set to unity. Figures 1 and 2 show, in box-plot form, respectively the variation in the 100-yr RMSE and BIAS values for different sets of catchment descriptors used in Eq. (1). In Table (2), the corresponding mean variation of $RMSE_{100}$ and $RMSE_{50}$ values, for different sets of pooling variables, is summarised. It shows that the numerical measures of effectiveness vary by very little between rows. The set of two variables, lnAREA and lnSAAR, and the set of the single variable lnAREA performed best in terms of providing the lowest $RMSE_{100}$ values. In terms of $RMSE_{50}$, the set consisting of lnAREA and lnSAAR comes second best to the set consisting of lnAREA on its own. Overall, the set of variables comprised of lnAREA and lnSAAR may be considered as being the most suitable set of pooling variables for Irish conditions. However, if there is also

HESSD
7, 5099–5130, 2010

a desire to incorporate another physical catchment effect then the BFI could be included with these two. While inclusion of just one or two catchment descriptors may indeed be best, there is an intuitive attraction in also representing some descriptor of catchment response even at the cost of a small apparent loss in effectiveness. An extension to this investigation with varying values of weights W_k in Eq. (1) was also done, particularly for the set of variables of lnAREA, lnSAAR and BFI but the results of all variations examined are not reported in detail here. It was found, by trial and error, that the weights 1.5, 1.0 and 0.1 for lnAREA, lnSAAR and BFI, respectively gave $\text{RMSE}_{100}=15.22$ and $\text{RMSE}_{50}=12.81$ which offer small improvements on the $W_k=1.0$ values used in the calculations for the set of variables of lnAREA, lnSAAR and BFI.

3 Procedure for examination of homogeneity

The procedure which is applied in the study is as follows

1. The gauging stations in the subject site's pooling group are identified using d_{ij} values obtained by the following equation having a minimum of 500 station years of data in the pooling group and satisfying the 5T rule for the 100 yr quantile.

$$d_{ij} = \sqrt{1.5\left(\frac{\text{lnA}_i - \text{lnA}_j}{\sigma_{\text{lnA}}}\right)^2 + \left(\frac{\text{lnSAAR}_i - \text{lnSAAR}_j}{\sigma_{\text{lnSAAR}}}\right)^2 + 0.1\left(\frac{\text{BFI}_i - \text{BFI}_j}{\sigma_{\text{BFI}}}\right)^2} \quad (6)$$

   The weights 1.5, 1.0 and 0.1 are those reported in the last paragraph.

2. The t_2 is obtained for each site in the pooling group and the average, without weights, of these is calculated to represent the pooled average t_2 (t_2^R).

3. Random samples are drawn from an EV1 distribution (or a GEV distribution with a selected shape parameter value) using the t_2^R as the population value to construct a 95% confidence interval for t_2^R. The sample size is taken as being equal

HESSD
7, 5099–5130, 2010

S. Das and C. Cunnane

to the average record length of the observed historical record at the gauging sites and the parameter values are estimated from the value of the t_2^R. The 95% confidence interval is constructed assuming that the samples t_2 are normally distributed.

4. The number of stations in the selected pooling group whose t_2 values fall outside the confidence interval (the attribute termed here as m) is counted and reported. It is also noted whether the t_2 of the subject site is outside the confidence limits (CL).

3.1 Analysis

The procedure described above is applied for each of the 85 stations. Each station had its own unique pooling group. The sample values of t_2 for the stations in the group, t_2^R and the CL about t_2^R are displayed in Fig. 3 for five stations. The GEV is used as the parent distribution to construct confidence limits. Three population shape parameters, $k=-0.05$, $k=0.0$ (EV1) and $k=0.03$, are selected in this context which correspond to L-skewness ≈0.21, 0.17 and 0.15, respectively, this being the range relevant for Ireland. The summary statistics of the procedure are given in tabular form in Table 3. In addition to that the heterogeneity measures, H1 and H2, described in Appendix A , for each group is calculated and a summary of these measures is reported in Table 4.

The following observations and findings are obtained from the analysis.

1. Table 3 lists how many stations fall into the categories of zero value outside the CL, one value outside the CL, 2 or 3 values outside the CL or more than 3 outside the CL. In all, for the case of EV1, only one station (1%) was in the first category while 52% of stations were in the latter category.

2. From Table 3, it is seen that as the shape parameter increases from $k=-0.05$ to $+0.03$ the number of cases where $m>3$ increases from 33 to 47.

HESSD
7, 5099–5130, 2010

3. In 27 groups (32% of groups) the t_2 of the subject site was outside the CL for the case of EV1. The corresponding numbers for the case of negative shaped GEV and for the case of positive shaped GEV are 27 and 28, respectively. All the 27 stations of the EV1 case were also in the later cases.

4. Table 4 summarises the results of H1 and H2 for the 85 pooling groups. 22% of groups have a H1 value lower than 4.0. The percentage increases to 86% when the same criterion is set for H2 and that is very similar to what was found for the UK pooling groups (FEH, 1999, p. 176).

5. The average range of t_2 for the 85 pooling groups was 0.11 with a minimum value of 0.06 and a maximum value of 0.18. Figure 4 shows a plot between H1 values and ranges of L-CV values for the 85 groups. The plot shows an upward trend, implying that a high H1 value can be expected for a high range of t_2 values in a pooling group, which can be expected in the absence of homogeneity. A similar plot is drawn for H2 in Fig. 5, showing no obvious trend, implying that a low H2 value may be obtained for a pooling group which is in fact a heterogeneous group.

6. Figure 6 shows a plot between H1 and *m*. Different values of H1 occur for a particular *m* value and that is reasonable as the memberships of the groups in those cases are different even though they may have some overlap. However, the average values, marked by triangles in the plot, show an increase of H1 with *m*, i.e. the higher the number of t_2 values of group members outside the CL, the higher the value of H1 that can be expected. If we consider a H1 value less than 4.0 as a good criterion for testing homogeneity, then in this approach we should not allow more than 2 values of t_2 to fall outside the confidence limits.

7. Figure 7 shows a plot between H1 and $d_{ij,\max}$ of the pooling groups. The $d_{ij,\max}$ is defined here as the distance associated with the group member which just qualified as a member of the pooling group. The plot shows an upward trend to some extent, implying that a low H1 value can be expected for a low $d_{ij,\max}$

value, which is an implicit assumption of a ROI pooling scheme. However in many cases, low $d_{ij,\max}$ values, even those below 1.0, can lead to a high value of H1 suggesting that the assumption may not always be true particularly for Irish conditions. A similar plot is drawn in Fig. 8 between $d_{ij,\max}$ and m. The plot leads to a similar conclusion to that for Fig. 7. While a low value of $d_{ij,\max}$ is desirable, it is noted that even low values of $d_{ij,\max}$ can occur where a significant number of group members' t_2 values falls outside the CL.

4 Investigation of selected heterogeneous pooling groups

The investigation has been carried out on those cases where the pooling groups are heterogeneous and in which the t_2 of the subject site lies outside the confidence limits. The investigation mainly focuses on identifying any inappropriateness among group members that would cause the pooling groups to be heterogeneous. In this context, FEH (1999, 3, Fig. 16.9) documented a detailed review system, providing an example. It mainly considers two attributes: 1) whether the subject site has any special qualities that need to be taken into account and 2) whether any of the pooled sites has catchment descriptors that are particularly different from those of the subject site.

Sites in the pooling group can be investigated using several characteristics including at-site flood statistics and catchment descriptors. Statistics in a pooling group such as discordancy measure (Hosking and Wallis, 1997) and the distance measure (d_{ij}) can also be used to investigate sites in the pooling group. In this part of the study, four catchment descriptors, namely, catchment size (AREA), wetness (SAAR), soils (BFI), lakes and reservoirs (FARL); and the distance measure (d_{ij}) are taken into account in the investigation process. The first three of the catchment descriptors were already used for initial selection of sites for a pooling group. In the investigation procedure, sites are reviewed with the help of Box-plots and a summary table and in some cases, with the help of an "examination of homogeneity" chart. Four Box-plots of catchment descriptors, such as AREA, SAAR, BFI and FARL, are constructed to show the subject site in the context of the pooling group. For each of these catchment descriptors,

HESSD
7, 5099–5130, 2010

the placement of values for sites in the pooling group is displayed against a backdrop of the relative frequency of the 85 sites considered in this study. This facilitates the identification of any particularly inappropriate sites. In the summary table, statistical properties such as t_2, t_3 and d_{ij} values of sites in a group are listed. The investigation procedure for pooling groups of station no 6031 is described in detail as it serves as an example.

4.0.1 An example: Station no 6031 on the River Flurry

There are 17 sites in the pooling group of which eight, including the subject site, have values which fall outside the CL, thus indicating a strongly heterogeneous group. The heterogeneity measures H1 and H2 for the group are 7.66 and 2.82, respectively. The examination of Box-plots in Fig. 9 reveals the catchment area of the subject site is small (46.2 km^2) and it is very near to the 5 percentile mark on the Box-plot of AREA. The site is not positioned at the centre of the group of gauged catchments in the pooling group. There are 5 sites on the left of the subject site and there are as many as 11 sites on the right. The attribute certainly includes some sites that have large catchment area compared to the subject site. This may lead to d_{ij} values exceeding the value 1.0 in several cases. The d_{ij} values for the last three sites are around 1.3 and these sites are among the seven other sites that fall outside the CL. The examination of the summary table on the right hand side of Fig. 9 shows that the subject site has large values of both t_2 and t_3 and that these are the largest among the group members. Hence, the conclusion can be drawn here that the pooling group in its present structure may not be ideal for that subject site 6031. Leaving out some sites at the bottom of the table might be considered in this context. The large number of sites, 17, in the pooling group is also a possible contributor to heterogeneity.

In case of the remaining 26 groups of such are investigated but the outcomes of the investigation procedure are not reported here. It is found that in many cases, special attributes of the subject site contributed to the degree of heterogeneity of the pooling groups.

HESSD
7, 5099–5130, 2010

5 Conclusions

In the context of ROI pooling group based flood frequency estimation procedure, the most suitable form of distance measure d_{ij} for Irish conditions was sought. The ROI method with the suitably identified distance measure, Eq. (6), was used to form pooling groups for the subject sites. A simple graphical approach of examining homogeneity of the pooling groups was presented. The graphical approach compared the sampling variability of pooled estimates of L-CV with the L-CV of pooling group members. The approach also allowed the location of L-CV of the subject site to be viewed in the context of pooling group members, which is important in the case of site specific pooling group. Most of the Irish pooling groups exhibited a degree of heterogeneity among the group members. A graphical approach of reviewing a heterogeneous pooling group was also presented in this context. The following conclusions were obtained from the above studies:

1. It was found that the distance measure d_{ij} could be satisfactorily defined in terms of lnAREA and lnSAAR but if there is a desire to incorporate another physical catchment effect then the BFI could be included with these two.

2. A visual approach for the identification of the homogeneity of ROI pooling groups has been presented. The results are compared with the heterogeneity measures H1 and H2, obtained for those groups. Overall the results show that even with a carefully considered ROI procedure, such as using distance measure of Eq. (6), it is not certain that perfectly homogeneous pooling groups are identified. As a compromise it is recommended that a group containing more than 2 values of L-CV outside the 95% confidence limits of that variable should not be considered homogeneous.

3. A thorough investigation on 27 heterogeneous pooling groups has been carried out. In many cases, special attributes of the subject site contributed to the degree of heterogeneity of the pooling groups. It is deemed necessary that the the

HESSD
7, 5099–5130, 2010
Homogeneity test
S. Das and C. Cunnane

subject site to be positioned near the centre of the group of gauging sites, on the respective catchment descriptor axes, to which it is hydrologically similar; but in some cases the fulfillment of that condition does not guarantee that the pooling group is homogeneous.

Appendix A

Heterogeneity test measures

The heterogeneity test measures proposed by Hosking and Wallis (1997) are based on 1) L-CV alone (the H1 statistic) and 2) L-CV and L-skewness jointly (the H2 statistic). These tests measures the sample variability of the L-moment ratios among the samples in the pooling group and compare it to the variation that would be expected in a homogeneous pooling group. The sample variability of the L-moment ratios is measured as the standard deviation of the at-site sample L-moment ratios weighted proportionally to the sites' respective record lengths. The measure of the sample variability based on L-CV alone, i.e. V_1, and L-CV & L-skewness jointly, i.e. V_2, are defined as

$$V_1 = \left[\sum_{i=1}^{M} n_i \left(t_2^i - t_2^R \right)^2 \Big/ \sum_{i=1}^{M} n_i \right]^{1/2} \quad \text{(A1)}$$

$$V_2 = \sum_{i=1}^{M} n_i \left[\left(t_2^i - t_2^R \right)^2 + \left(t_3^i - t_3^R \right)^2 \right]^{1/2} \Big/ \sum_{i=1}^{M} n_i \quad \text{(A2)}$$

where t_2^R and t_3^R are the group average of L-CV and L-skewness, respectively; t_2^i , t_3^i and n_i are the values of L-CV, L-skewness and the sample size for site i and M is the number of sites in the pooling group.

HESSD
7, 5099–5130, 2010

Simulation is used to establish what "would be expected" of a homogeneous group. Some 500 homogeneous groups are generated using a four-parameter kappa distribution with L-moment ratio values equal to t_2^R, t_3^R, t_4^R and the at-site mean, $L1=1$, in order to obtain the expected mean value, μ_{V_j}, and the standard deviation, σ_{V_j}, of the variability measures for a homogeneous group.

The heterogeneity measures H_j are then estimated using the expression below.

$$H_j = \frac{\left(V_j - \mu_{V_j}\right)}{\sigma_{V_j}}, \quad \text{for } j = 1,2 \tag{A3}$$

Hosking and Wallis (1997) recommended using the H1 statistic over the H2 statistic as they found that the heterogeneity measure based on V_1 has better power to discriminate between homogeneous and heterogeneous regions. They suggested that a region is considered to be "acceptably homogeneous" if $H1<1$, "possibly heterogeneous" if $1<H1<2$, and "definitely heterogeneous" if $H1>2$. However, FEH (1999) opted for the H2 statistic as the heterogeneity measure for testing the homogeneity of pooling groups as both the L-CV and L-skewness are required for fitting pooled growth curves with a Generalised Logistic (GLO) or Generalised Extreme Value distribution (GEV). They revised the heterogeneity criteria based on the H2 statistics, suggesting that if $2<H2<4$, a region could be considered as heterogeneous whereas if $H2>4$ it could be considered as strongly heterogeneous.

Appendix B

Estimation of growth factor

The dimensionless L-moment ratios t_2 and t_3 are calculated for each site. Pooled L-moment ratios for the target site, i, are then computed using the following equation:

$$t^{(i)R} = \frac{\sum_{j=1}^{M} w_{ij} t^{(j)}}{\sum_{j=1}^{M} w_{ij}} \tag{B1}$$

where $t^{(j)}$ is the L-moment ratio (either t_2 or t_3) for the jth most similar site and w_{ij} is a weighing term, taken here as 1.

The Generalised Extreme Value (GEV) has been selected as the pooled distribution function. The values t_2^R, t_3^R are equated to expressions for these quantities written in terms of the distribution's unknown parameters (expressed in dimensionless form) and the resulting equations are solved for the unknown parameter values. The dimensionless GEV growth curve ($X_T = Q_T / Q_{\text{med}}$) is defined by two parameters k and β:

$$X_T = 1 + \frac{\beta}{k}\left((\ln 2)^k - \left(-\ln \frac{T}{T-1} \right)^k \right) \tag{B2}$$

where T is the return period.

The two parameters k and β are estimated from the sample L-CV, t_2, and sample L-skewness, t_3, as follows (Hosking and Wallis, 1997)

$$k = 7.8590c + 2.9554c^2 \tag{B3}$$

in which

$$c = \frac{2}{3+t_3} - \frac{\ln 2}{\ln 3} \tag{B4}$$

$$\beta = \frac{kt_2}{t_2\left(\Gamma(1+k)-(\ln 2)^k\right)+\Gamma(1+k)\left(1-2^{-k}\right)} \tag{B5}$$

where Γ denotes the complete gamma function.

Acknowledgements. The authors would like to acknowledge the financial support made available by the Irish Office of Public Works.

References

Burn, D. H.: Evaluation of regional flood frequency-analysis with a region of influence approach, Water Resour. Res., 26, 2257–2265, 1990. 5103

Castellarin, A., Burn, D. H., and Brath, A.: Assessing the effectiveness of hydrological similarity measures for flood frequency analysis, J. Hydrol., 241, 270–285, 2001. 5104

Chowdhury, J. U., Stedinger, J. R., and Lu, L. H.: Goodness-of-fit tests for regional generalized extreme value flood distributions, Water Resour. Res., 27, 1765–1776, 1991. 5100, 5101

Dalrymple, T.: Flood frequency analysis, Water Supply Paper 1543 A, US Geol. Surv., 1960. 5100, 5101, 5102

FEH: Flood Estimation Handbook, vol. 1–5, Institute of Hydrology, Wallingford, 1999. 5100, 5101, 5103, 5109, 5110, 5114

Fill, H. D. and Stedinger, J. R.: Homogeneity tests based upon Gumbel distribution and a critical-appraisal of Dalrymple test, J. Hydrol., 166, 81–105, 1995. 5100, 5101

FSR: Flood Studies Report, vol. 1, Nat. Environ. Res. Council, London, 1975. 5101

FSU: Irish Flood Studies Update – Work Package 2.3, Office of Public Works, Dublin, 2009. 5104

Gaál, L., Kyselý, J., and Szolgay, J.: Region-of-influence approach to a frequency analysis of heavy precipitation in Slovakia, Hydrol. Earth Syst. Sci., 12, 825–839, doi:10.5194/hess-12-825-2008, 2008. 5104

Hosking, J. and Wallis, J. R.: Regional Frequency Analysis, Cambridge University Press, Cambridge, 1997. 5100, 5101, 5102, 5104, 5110, 5113, 5114, 5115

Hosking, J. R. M. and Wallis, J. R.: Some statistics useful in regional frequency-analysis, Water Resour. Res., 29, 271–281, 1993. 5100, 5102

Lettenmaier, D. P., Wallis, J. R., and Wood, E. F.: Effect of regional heterogeneity on flood frequency estimation, Water Resour. Res., 23, 313–323, 1987. 5101

Lu, L. H. and Stedinger, J. R.: Sampling variance of normalized GEV PWM quantile estimators and a regional homogeneity test, J. Hydrol., 138, 223–245, 1992. 5100, 5102

Stedinger, J. R. and Lu, L. H.: Appraisal of regional and index flood quantile estimators, Stoch. Hydrol. Hydraul., 9, 49–75, 1995. 5101

Viglione, A., Laio, F., and Claps, P.: A comparison of homogeneity tests for regional frequency analysis, Water Resour. Res., 43, W03428, doi:10.1029/2006WR005095, 2007. 5100, 5102

Wiltshire, S. E.: Identification of homogeneous regions for flood frequency-analysis, J. Hydrol., 84, 287–302, 1986. 5100, 5101

HESSD

7, 5099–5130, 2010

Table 1. Summary of AMF data sets used in the study.

Number of stations	85
Shortest record length	17
Longest record length	55
Mean record length	36.5
Number of AMF events	3213

HESSD

7, 5099–5130, 2010

Table 2. Variation in the mean RMSE corresponding to $T=100$ and 50 for different sets of pooling variables.

Variables used in model	RMSE100%	RMSE50%
lnAREA (lnA)	15.13	12.47
lnAREA,lnSAAR (lnA+lnS)	15.11	12.77
lnAREA,lnSAAR,BFI (lnA+lnS+B)	15.52	13.22
lnAREA,lnSAAR,BFI,FARL (lnA+lnS+B+F)	15.57	13.20
lnSAAR (lnS)	15.27	13.23
BFI (B)	15.97	13.83
lnAREA,BFI (lnA+B)	16.21	13.44
lnAREA,lnSAAR,FARL (lnA+lnS+F)	15.54	12.78

HESSD
7, 5099–5130, 2010

Table 3. Summary of events outside the confidence limits for 85 pooling groups.

events outside the CL (m)	GEV (k=−0.05)	EV1 (k=0)	GEV (k=+0.03)
	No. of groups (%)	No. of groups (%)	No. of groups (%)
0	2 (2)	1 (1)	0 (0)
1	13 (15)	7 (8)	7 (8)
2 <= 3	36 (42)	33 (39)	31 (36)
> 3	34 (40)	44 (52)	47 (55)

HESSD

7, 5099–5130, 2010

S. Das and C. Cunnane

Table 4. Summary of heterogeneity measure, H1 and H2 for 85 pooling groups.

Heterogeneity measure	% of groups with heterogeneity<2	% of groups with heterogeneity<4
H1	5	22
H2	38	86

HESSD
7, 5099–5130, 2010

Homogeneity test
S. Das and C. Cunnane

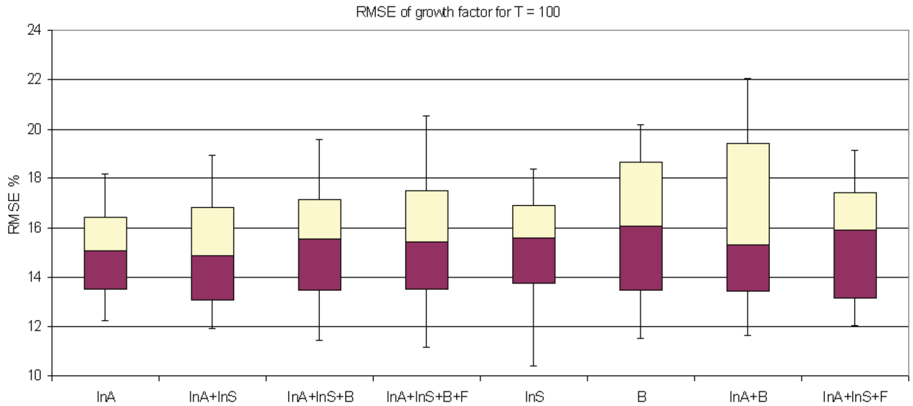


Fig. 1. Box-plot of RMSE of growth factors corresponding to 100 yr return periods for different sets of catchment descriptors used in defining the distance measure d_{ij}. Each Box-plot gives the percentiles for the frequencies 0.05, 0.25, 0.5, 0.75, 0.95.

HESSD
7, 5099–5130, 2010
Homogeneity test
S. Das and C. Cunnane

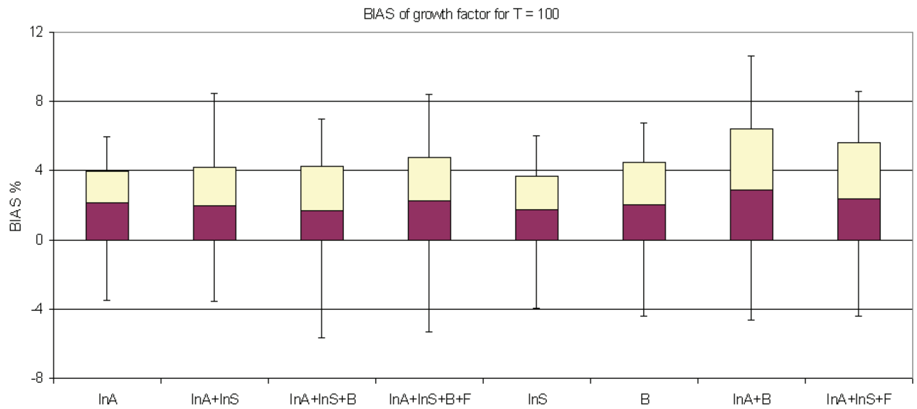


Fig. 2. Values of BIAS for the analysis summarised in Fig. 1.

HESSD
7, 5099–5130, 2010

Homogeneity test

S. Das and C. Cunnane

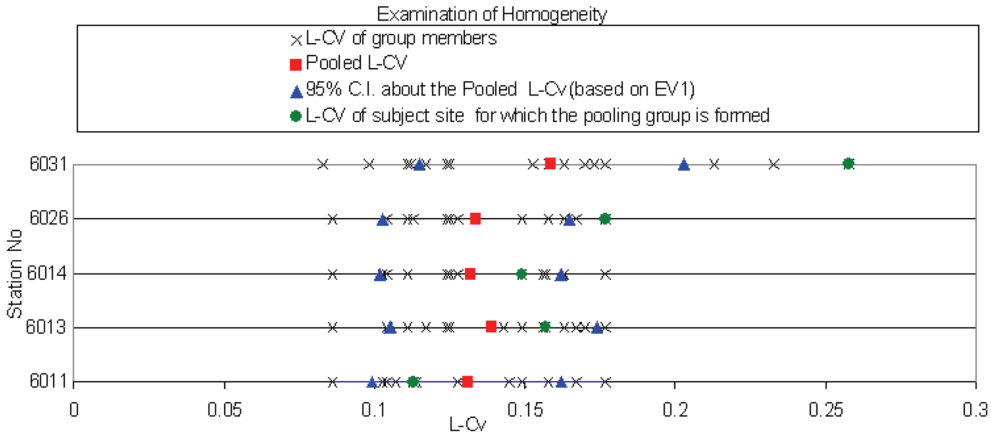


Fig. 3. Examination of homogeneity.

HESSD
7, 5099–5130, 2010

Homogeneity test

S. Das and C. Cunnane

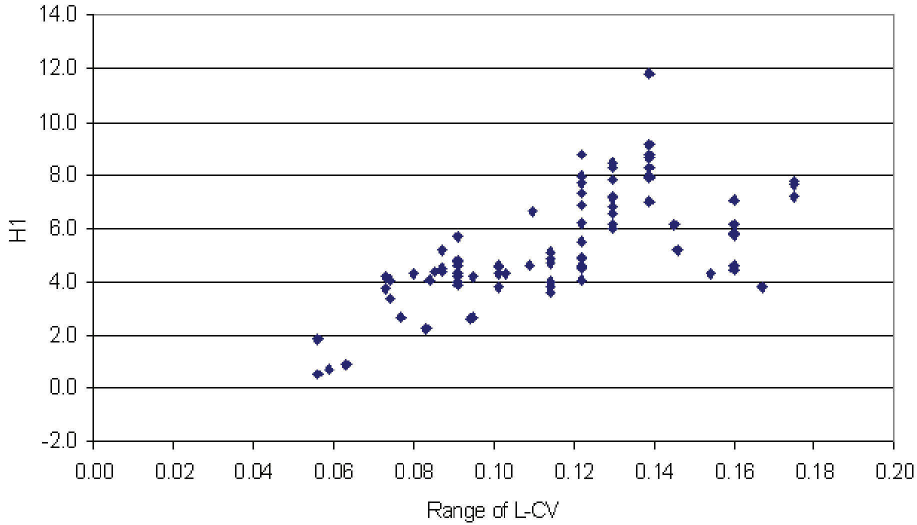


Fig. 4. H1 plotted versus range of L-CV. Each point represents a pooling group.

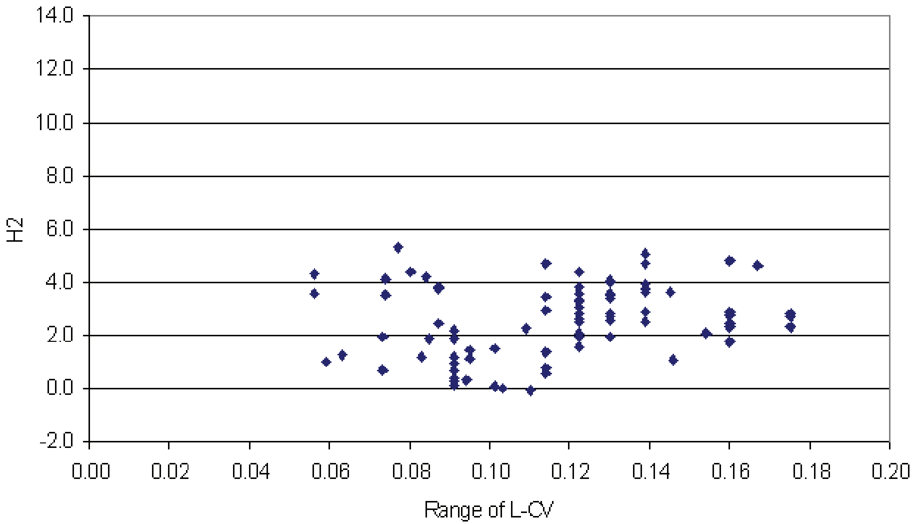


Fig. 5. H2 plotted versus range of L-CV. Each point represents a pooling group.

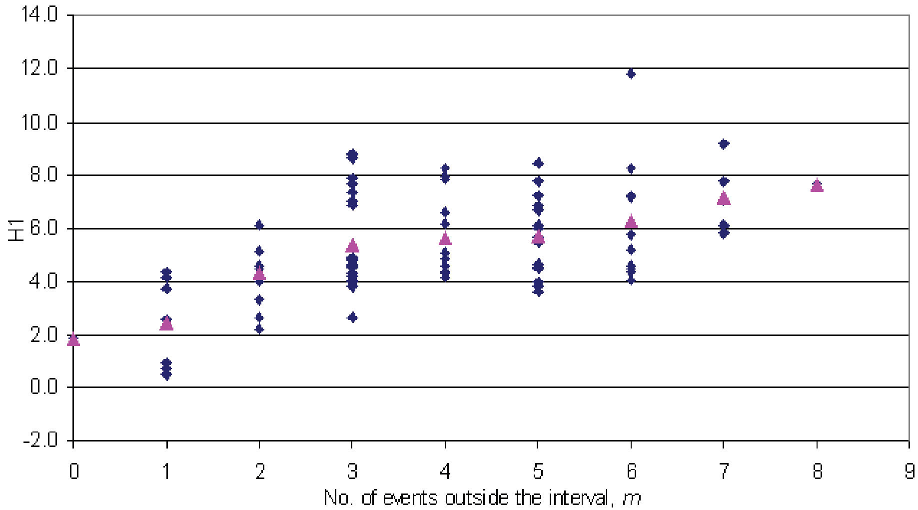


Fig. 6. H1 plotted versus *m*. Each point represents a pooling group.

HESSD
7, 5099–5130, 2010

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◀| |▶
◀ ▶
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion

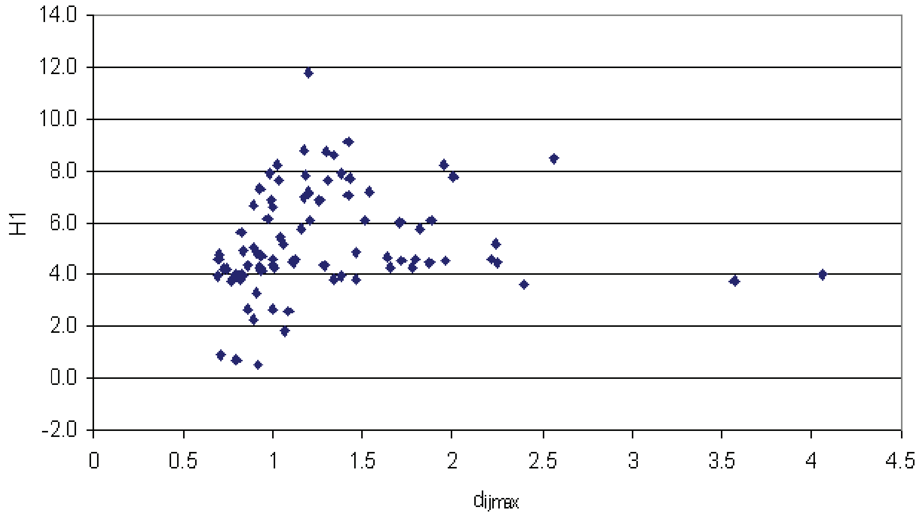


Fig. 7. H1 plotted versus $d_{ij,\max}$. Each point represents a pooling group.

HESSD
7, 5099–5130, 2010

Homogeneity test

S. Das and C. Cunnane

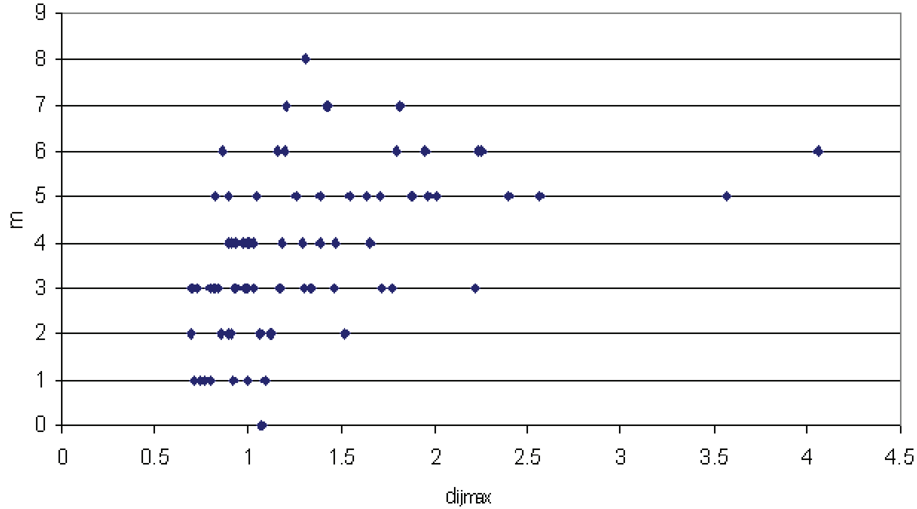
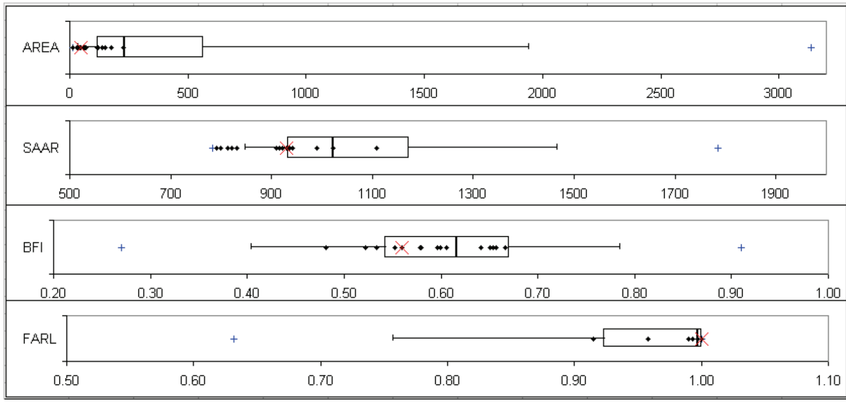


Fig. 8. *m* plotted versus $d_{ij,\max}$. Each point represents a pooling group.

HESSD
7, 5099–5130, 2010

Homogeneity test

S. Das and C. Cunnane



St.No	RL	t2	t3	dij
6031	18	0.258	0.392	0
24022	20	0.233	0.123	0.277
26022	33	0.173	0.091	0.296
36031	30	0.098	0.365	0.76
25040	20	0.153	0.203	0.809
16001	33	0.124	0.012	0.938
25023	52	0.163	0.142	0.988
8002	20	0.112	0.271	1.021
7006	19	0.117	-0.205	1.068
14009	25	0.125	0.136	1.092
25027	43	0.158	0.046	1.144
10021	24	0.213	0.186	1.146
6026	46	0.177	0.24	1.192
14007	25	0.17	0.219	1.234
10022	18	0.233	0.057	1.274
16004	48	0.111	0.069	1.276
29004	32	0.083	0.144	1.306

Fig. 9. Four Box-plots and a summary table for investigating a pooling group. The subject site is marked with an ×. Small dots denote sites included in the pooling group. The underlying distribution of each catchment descriptor is shown in the Box-plots. Each Box-plot gives the minimum and the maximum value (+) and percentiles for the frequencies 0.05, 0.25, 0.5, 0.75, 0.95. The summary table lists record length, t_2, t_3 and d_{ij} values for a 100-yr pooling group for subject station 6031.

HESSD

7, 5099–5130, 2010

Homogeneity test

S. Das and C. Cunnane

Title Page
Abstract | Introduction
Conclusions | References
Tables | Figures
◂◂ | ▸▸
◂ | ▸
Back | Close
Full Screen / Esc
Printer-friendly Version
Interactive Discussion