Revista Colombiana de Estadística

No.33 y No.34

REGIONAL FLOW ANALYSIS FOR THE VALLE DEL CAUCA REGION IN COLOMBIA

C. A. GONZÁLEZ, PACHECO. P. N AND FEYEN J.

ABSTRACT. Flow frequency analysis of high and low flows is carry out by using the regional Probability Weighted Moments (PWM) algorithm. The distributions fitted to the region are: General extreme value (GEV), Wakeby, Extreme value type 1 (EV1) or Gumbel and Weibull (two or three parameters) distributions. In order to select the distribution appropite for each series four test of goodness of fit were carry out, manely, chi square, 'DMAX ' goodness of fit, probability plot criteria and a graphical test base on montecarlo simulation. The selection of the distribution is made to depend on the performance of the distributions under tests over all test. The distributions selected for annual maximum flow series is the EV1 distribution while the GEV distribution is selected for the annual minimum flow series.

INTRODUCTION

General information

Colombia is located on the Equator in the North-west corner of South America, approximately between 4° 15 ' South latitude and 12° 20 ' North latitude and between 66° 50 ' and 79° 2 ' West longitude.

The Cauca river is the main river drainage basin of the study region, which from an economic point of view is the second most important river in the country. It raises in the Andes mountain range of Southern Colombia and flows Northwards into the Atlantic Ocean. The study region is crossed from South to North by the Western mountain range, with a mean altitude of 2000 m above msl. This mountain range divides the water which drains to the Pacific ocean, and that which drains Eastwards to the Cauca river at 950m above mls approximately. Parallel to that range and to the East of it is the Central mountain range with a mean altitude of 3000m above msl approximately, which separate the catchments of Cauca and Magdalena rivers.

The climate of the region is classified as tropical due to its geographical position. This type of climate is characterized by its temperature whose average variation is not greater than 5° C. However, the climate is strongly influenced by the orography of the region which affects the mean temperature according to altitude above mean sea level. Another factor which affects the climate together with the orography the atmospheric circulation in the form of winds, of the air that is present in the valley region.

From the monthly maps of area distribution of rainfall obtained from González D. (1984), it is observed that there are two periods of drought and two of rainfall each year. The drought periods are from January to March and from June to August, while the rainfall periods are in April and May and September to December . It is important to remark that despite the generally similar patterns of all maps , there exists significant variation between the different stations in regard to the distribution of rainfall during the year, reflecting climatic variations in the region.

Stations selected

Table 1 shows the names of the hydrometric stations situated on the Cauca River and its tributaries which were selected for the study. Table 2 shows the period for which records were available. Table 2 shows that records were kept from most of the tributaries of the Cauca River, from 1973 up to 1984. However, most of the stations

Table 1. Name and Location of the hydrometric stations

Número	Nombre Estación	Rio	Longitud grados.min	Latitud grados.min	Ciudad	Dpto	Altura (msnm)	Area (km2)	Instrumento de registro	Fecha de instalación mes-año
1	Salvajina	Cauca	76,43	2,57	B/Aires	Cauca	1029	3652	c.s.r*	VIII-1946
2	La Balsa	Cauca	76,36	3,05	B/Aires	Cauca	986	5156	c.s.r	X-1976
3	La Bolsa	Cauca	76,3	3,13	Santander	Cauca	964	6110	c.s.r	VII-1967
4	Hormiguero	Cauca	76,29	3,18	Cali	Valle	956	8018	d.s.r	X-1961
5	Juanchito	Cauca	76,29	3,27	Candelaria	Valle	949	8673	c.s.r	I-1934
6	Pto. Isaacs	Cauca	76,29	3,39	Yumbo	Valle	945	9079	d.s.r	IV-1955
7	Media Canoa	Cauca	76,21	3,53	Yotoco	Valle	934	12176	c.s.r	III-1961
8	Rio Frio	Cauca	76,16	4,08	Rio Frio	Valle	924	13176	c.s.r	X-1958
9	Guayabal	Cauca	76,06	4,24	Zarzal	Valle	912	16036	d.s.r**	III-1955
10	La Victoria	Cauca	76,02	4,31	La Victoria	Valle	907	16284	c.s.r	X-1958
11	Anacaro	Cauca	75,58	4,37	Cartago	Valle	902	17234	c.s.r	III-1955
12	La Luisa	Claro	76,36	3,18	Jamundi	Valle	1034	85	c.s.r	IX-1945
13	Ortigal	Desbaratado	76,21	3,17	Florida	Valle	988	105	d.s.r	IX-1971
14	Buchitolo	Fraile	76,21	3,23	Candelaria	Valle	981	263	c.s.r	VII-1961
15	El Vergel	Guadalajara	76,16	3,53	Buga	Valle	1090	131	c.s.r	VIII-1971
16	Los Bueyes	Guengue	76,27	3,13	Corinto	Cauca	1150	93	d.s.r	III-1971
17	Potrerito	Jamundi	76,36	3,15	Jamundi	Valle	1010	69	c.s.r	I-1946
18	La Sorpresa	La Paila	76,04	4,2	Zarzal	Valle	950	316	c.s.r	IX-1971
19	Cabito	La Quebrada	76,24	3,04	Caloto	Cauca	1060	60	d.s.r	II-1971
20	Lomitas	La Teta	76,35	3,03	B/Aires	Cauca	1022	128	c.s.r	II-191
21	Caicedonia	La Vieja	75,53	4,25	Caicedonia	Valle	1080	2038	c.s.r	IX-1971
22	Cartago	La Vieja	75,53	4,45	Cartago	Valle	914	3285	d.s.r	X-1945
23	Media Canoa	Media Canoa	76,23	3,54	Yotoco	Valle	943	73	d.s.r	IX-1971
24	Santa Librada	Morales	76,14	4,1	Tulua	Valle	940	170	d.s.r	X-1972
25	Abajo	Ovejas	76,36	2,52	B/Aires	Cauca	1263	607	d.s.r	V-1964
26	Pto. Tejada	Palo	76,26	3,14	Pto. Tejada	Cauca	966	1656	c.s.r	IX-1954
27	Pichinde	Pichinde	76,37	3,26	Cali	Valle	1540	52	c.s.r	IX-1968
28	Pte.Ferrocarril	Quinamayo	76,31	3,02	Santander	Cauca	1006	160	c.s.r	I-1970
29	Timba	Timba	76,37	3,07	B/Aires	Cauca	997	401	c.s.r	VIII-1946
30	Mateguadua	Tulua	76,1	4,01	Tulua	Valle	1115	732	c.s.r	IX-1945

**c.s.r: Limnígrafo
*d.s.r: Limnímetro

Table 2. Records for daily flow available

Número	Nombre Estación	Río	1934	1935	1936	1937	1938	1939	1940	1941	1942	1943	1944	1945	1946	1947	1948	1949	1950	1951	1952	1953	1954	1955	1956	1957	1958	
1	Salvajina	Cauca	+	+	+	+	+	+	+	+	+	+	+	+	10	*	*	*	*	3_10	*	*	*	*	*	*	*	
2	La Balsa	Cauca	+	+	+	+	+	+	+	+	+	+	+	+	1	*	+	*	*	3_10	*	*	*	*	*	*	*	
3	La Balsa	Cauca	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
4	Hormiguero	Cauca	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
5	Juanchito	Cauca	1	8_	_	_	_	1	*	*	*	*	8_	10	*	*	*	8_	6	*	*	*	*	*	*	*	*	
6	Pte. Isaacs	Cauca	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
7	Mediacanoa	Cauca	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
8	Río Frío	Cauca	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	10	+	
9	Guayabal	Cauca	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	11	
10	La Victoria	Cauca	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	10	
11	Anacaro	Cauca	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
12	La Luisa	Claro	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1	*	*	*	*	*	*	*	*	
13	Ortigal	Desbaratado	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
14	Buchitolo	Fraile	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
15	El Vergel	Guadalajara	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
16	Los Bueyes	Guengue	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
17	Potrerito	Jamundí	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1	*	*	*	12_	_	_	_	_	10	+	
18	La Sorpresa	La Paila	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
19	Cabilo	La Quebrada	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
20	Lomitas	La Teta	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
21	Caicedonia	La Vieja	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
22	Cartago	La Vieja	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
23	Media Canoa	Media Canoa	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
24	Santa Librada	Morales	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
25	Abajo	Ovejas	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
26	Pte. Tolosa	Palo	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
27	Pichindé	Pichindé	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
28	Pte. Ferrocarril	Quinamayó	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	+	
29	Timba	Timba	+	+	+	+	+	+	+	+	+	+	+	+	8	*	*	*	*	*	*	*	*	*	*	7_	_	
30	Mateguadua	Tuluá	+	+	+	+	+	+	+	+	+	+	+	3	*	*	*	*	*	*	6_	_	_	_	_	_	_	

Número	Nombre Estación	Río	1959	1960	1961	1962	1963	1964	1965	1966	1967	1968	1969	1970	1971	1972	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982	1983	1984
1	Salvajina	Cauca	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	12	*
2	La Balsa	Cauca	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	12	*	*
3	La Balsa	Cauca	+	+	+	+	+	+	+	+	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	12	*	*
4	Hormiguero	Cauca	+	+	+	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	12	*	*
5	Juanchito	Cauca	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	12	*	*
6	Pte. Isaacs	Cauca	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1	*	*	*	*	*	*	*	*	12	*	*
7	Mediacanoa	Cauca	+	+	+	+	+	+	8	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	12	*	*
8	Río Frío	Cauca	+	11_	_2	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	12	*	*
9	Guayabal	Cauca	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	12	*	*
10	La Victoria	Cauca	*	7_	_10	*	10_12	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	12	*	*
11	Anacaro	Cauca	+	+	+	+	+	+	+	+	+	1	*	*	*	*	*	*	*	*	*	*	*	*	*	12	*	*
12	La Luisa	Claro	*	*	*	*	*	*	*	4_5	*	*	*	*	6_	_	_1	*	*	*	*	*	*	*	*	*	*	12
13	Ortigal	Desbaratado	+	+	+	+	+	+	+	+	+	+	+	+	+	4	*	*	*	*	*	*	*	*	*	*	*	12
14	Buchitolo	Fraile	+	+	6	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	12		
15	El Vergel	Guadalajara	+	+	+	+	+	+	+	+	+	+	+	+	+	2	*	*	*	*	*	*	*	*	*	*	*	3
16	Los Bueyes	Guengue	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1	*	*	*	*	*	*	*	*	*	12	
17	Potrerito	Jamundí	*	*	*	*	*	*	8_	_2	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	12
18	La Sorpresa	La Paila	+	+	+	+	+	+	+	+	+	+	+	+	+	1	*	*	*	*	*	*	*	*	*	*	*	12
19	Cabilo	La Quebrada	+	+	+	+	+	+	+	+	+	+	+	+	2	*	*	*	*	*	*	*	*	*	*	*	*	6
20	Lomitas	La Teta	+	+	+	+	+	+	+	+	+	+	+	+	+	1	*	*	*	*	*	*	*	*	*	*	*	12
21	Caicedonia	La Vieja	+	+	+	+	+	+	+	+	+	+	+	+	+	1	*	*	*	*	*	1	*	*	*	*	*	12
22	Cartago	La Vieja	+	+	+	+	+	+	+	+	+	+	+	+	+	1	*	*	*	*	*	*	*	*	*	*	*	12
23	Media Canoa	Media Canoa	+	+	+	+	+	+	+	+	+	+	+	+	+	1	*	*	*	*	*	*	*	*	*	*	*	12
24	Santa Librada	Morales	+	+	+	+	+	+	+	+	+	+	+	+	+	+	1	*	*	*	*	*	*	*	*	*	*	12
25	Abajo	Ovejas	+	+	+	+	+	6	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	12
26	Pte. Tolosa	Palo	+	+	+	+	+	+	+	1	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	*	12
27	Pichindé	Pichindé	+	+	+	+	+	+	+	+	+	+	+	+	+	+	2	*	*	*	*	*	*	*	*	*	*	12
28	Pte. Ferrocarril	Quinamayó	+	+	+	+	+	+	+	+	+	+	+	1	*	*	*	*	*	*	*	*	*	*	*	*	*	12
29	Timba	Timba	_	_	_	_	_	10	*	*	1_3	*	4_	_1	*	*	*	*	*	*	*	*	*	*	*	*	*	12
30	Mateguadua	Tuluá	_	_	_	_	_	_	_	_	5	*	*	*	2_	_3	*	*	*	*	*	*	*	*	*	*	*	11

*Registros Disponibles   +Sin Registro   _Período de Interrupción on the Cauca River have a longer recording period. As 1985 started to run the Salvajina dam which affects all the measurements downstream of the Cauca River. It was chosen a concurrent period of ten years from 1973 to 1982, being the longest period for which reliable concurrent information data are available. Given the effect of Salvajina dam on the floods and low flows of the Cauca River stations, the regional results obtained will be useful for application in the tributaries of the Cauca river.

Figure 1 illustrates the hydrometric stations in the region and the measurement instruments used. The instruments used are discrete and continuous stage recording.

Compilation of data

The series of extremes data were obtained from the Regional Corporation of Cauca River (C.V.C.). The annual maximum (Am) flow series is solely due to rainfall as in this region snowmelt is virtually absent.

The soil permeability varies significantly with the variation of soil types. For instance, it is known from direct observation that the soil type varies from clay to sand and in some places to stony strata. This variation is not only between catchments but also it is common to find this variation within the same catchment. It has undoubtedly a big influence on the catchment response to the rainfall input and hence on the characteristics of the series. However, the available geological information is still scattered and requires a further effort to assemble it into a more useful form that can be incorporated into a hydrological study.

In addition, the information on precipitation in the region, is presented in a general form to cover the whole region, averaging the rainfall patterns in the area which cannot be easily associated with any of the components catchments.

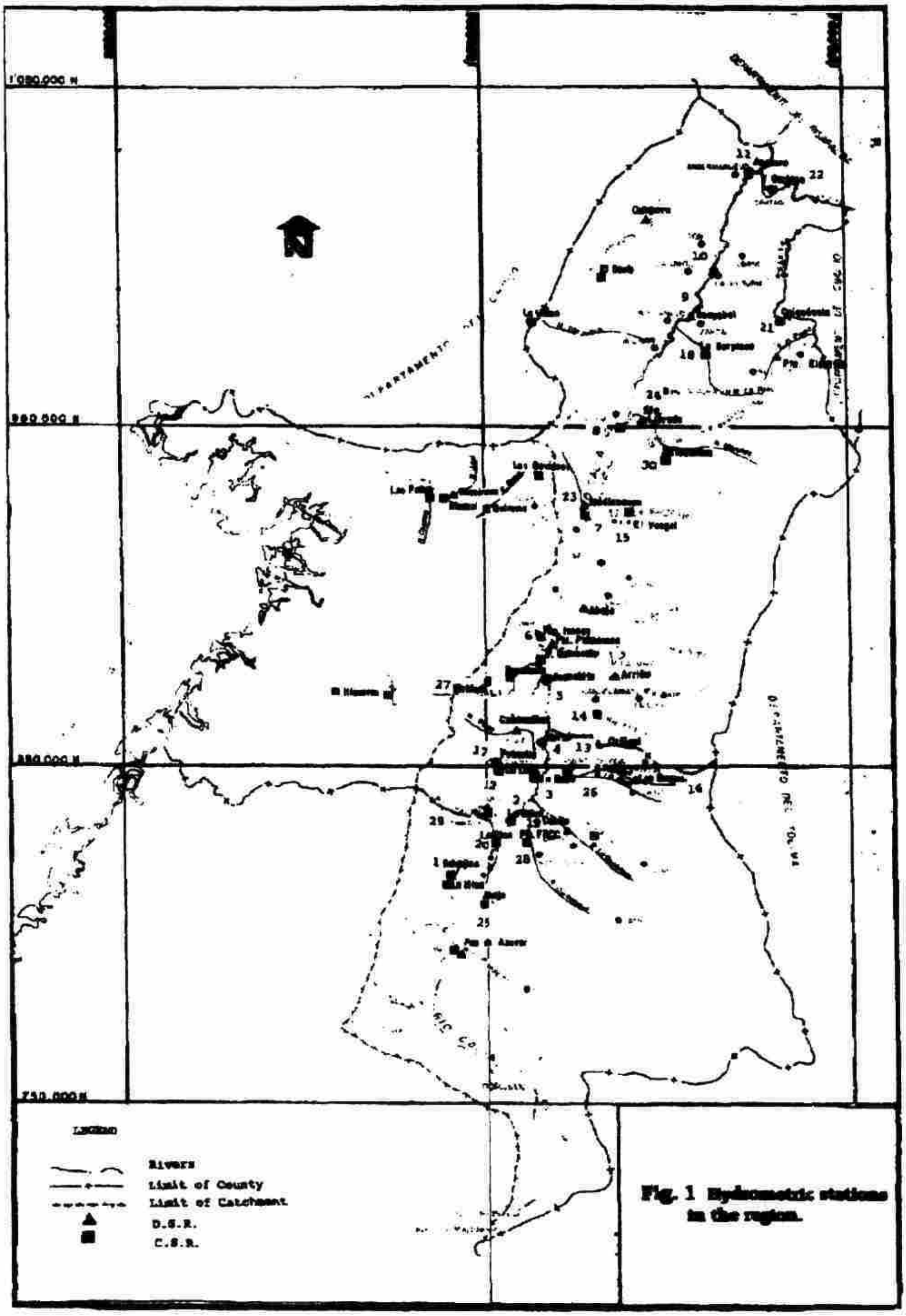


Fig. 1 Hydrometric stations in the region.

The minimum annual flows occurs between August and October. This can be explained by the fact that rivers are in recession since the beginning of the preceding dry period in June and because the period from September to October is usually too small to reserve these recession.

Tables 3 and 4 show the annual maximum and minimum flow series for equal concurrent lengths of record, for the thirty stations selected.

Preliminary Analyses

Basic statistics of the region, such as coefficients of variation and skewness and the variations on them, were compared with those of other regions for which studies have been produced, e.g. those carry out by Hosking et al.,(1985a), King (1985) and Wallis and Wood (1985). The statistics showed a region more heterogeneous in the annual maximum flow series and in the minimum flow series.

Tests of randomness of the data

In flood frequency the basic objective is the analysis of outcomes of real observations considered as a random sample. For this analysis, mathematical models are considered as the most condensed manner of expressing the information about random variables. The element of randomness is essential to the development of a statistical argument; in fact, it is the element that allows delineation of a population or collective from a description of which we can make probabilistic conclusions. Since hydrological situations can never be very well described, the question of independence, as that of randomness, is always problematical.

In order to check randomness to validate the use of the statistical inferences pro-

Table 3. Annual Maximum Flow Series

Número	Nombre Estación	Rio	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982
1	Salvajina	Cauca	552,00	1037,00	815,00	638,00	490,00	497,00	445,00	376,00	760,00	566,00
2	La Balsa	Cauca	900,00	1057,00	861,00	615,00	503,00	676,00	546,00	394,00	850,00	734,00
3	La Bolsa	Cauca	745,00	726,00	763,00	670,00	551,00	698,00	611,00	437,00	684,00	721,00
4	Hormigero	Cauca	902,00	886,00	981,00	811,00	605,00	817,00	850,00	463,00	769,00	890,00
5	Juanchito	Cauca	912,00	998,00	950,00	876,00	637,00	770,00	859,00	463,00	791,00	868,00
6	Pto. Isaacs	Cauca	917,00	973,00	1028,00	912,00	665,00	777,00	868,00	476,00	778,00	859,00
7	Media Canoa	Cauca	906,00	909,00	991,00	950,00	630,00	691,00	714,00	499,00	721,00	808,00
8	Rio Frio	Cauca	943,00	957,00	1033,00	1039,00	724,00	761,00	779,00	548,00	842,00	970,00
9	Guayabal	Cauca	1174,00	1163,00	1244,00	1203,00	817,00	902,00	900,00	548,00	997,00	972,00
10	La Victoria	Cauca	1130,00	1219,00	1324,00	1306,00	817,00	910,00	910,00	578,00	1000,00	975,00
11	Ancaro	Cauca	1199,00	1227,00	1276,00	1255,00	871,00	931,00	1020,00	594,00	1015,00	1066,00
12	La Luisa	Claro	71,20	49,60	71,20	42,00	79,50	59,70	76,80	50,10	34,00	89,90
13	Ortigal	Desbaratado	11,43	21,35	26,51	18,37	10,52	20,74	15,71	11,82	13,69	18,60
14	Buchitolo	Fraile	35,00	55,00	36,90	41,20	20,30	35,50	21,70	16,70	29,70	35,50
15	El Vergel	Guadalajara	46,00	122,10	79,90	45,20	27,20	57,70	49,30	57,40	230,70	52,40
16	Los Bueyes	Guengue	52,70	25,90	52,20	61,30	29,70	106,50	51,40	41,40	23,10	57,80
17	Potrerito	Jamundi	51,20	69,50	107,70	39,90	72,00	37,60	61,70	30,80	45,90	80,90
18	La Sorpresa	La Paila	148,81	142,26	122,89	79,02	94,50	183,00	138,20	57,84	164,70	130,70
19	Cabito	La Quebrada	3,60	58,60	55,70	27,30	41,60	31,10	23,70	36,20	47,10	44,30
20	Lomitas	La Teta	20,50	21,40	20,00	20,80	71,50	17,90	19,20	14,90	20,30	19,80
21	Caicedonia	La Vieja	667,00	496,00	717,00	474,00	28,00	350,00	423,00	202,00	492,00	483,00
22	Cartago	La Vieja	826,00	741,00	786,00	557,00	402,00	530,00	621,00	254,00	594,00	582,00
23	Media Canoa	Media Canoa	40,86	40,54	42,95	37,47	45,28	51,03	51,45	45,43	80,41	80,40
24	Santa Librada	Morales	24,10	27,20	20,90	20,90	20,80	22,20	20,90	22,80	19,90	19,90
25	Abajo	Ovejas	134,00	251,90	135,90	135,90	64,60	90,50	99,10	101,40	67,80	103,70
26	Pto. Tejada	Palo	284,90	263,20	285,50	237,40	159,00	290,60	251,30	150,40	207,80	264,90
27	Pichinde	Pichinde	20,00	79,90	31,40	22,50	17,10	16,00	17,40	11,70	38,30	29,60
28	Pte.Ferrocaril	Quinamayo	107,60	126,40	122,70	128,30	82,50	124,50	129,30	126,50	250,70	89,00
29	Timba	Timba	117,40	118,80	98,30	83,40	85,10	75,80	86,60	79,30	127,40	156,90
30	Meteguadua	Tulua	75,80	221,40	176,40	75,90	77,40	97,00	104,80	102,80	99,10	139,50

Table 4. Annual Minimum Flow Series

Número	Nombre Estación	Rio	1973	1974	1975	1976	1977	1978	1979	1980	1981	1982
1	Salvajina	Cauca	34,00	46,00	72,00	26,00	30,00	34,00	34,00	27,00	40,00	43,00
2	La Balsa	Cauca	47,00	82,00	86,00	40,00	46,00	44,00	55,00	40,00	50,00	64,00
3	La Bolsa	Cauca	51,00	71,00	125,00	49,00	52,00	53,00	61,00	44,00	66,00	67,00
4	Hormigero	Cauca	53,00	86,00	141,00	55,00	57,00	69,00	74,00	60,00	79,00	81,00
5	Juanchito	Cauca	59,00	91,00	148,00	59,00	61,00	71,00	77,00	59,00	79,00	81,00
6	Pto. Isaacs	Cauca	67,00	91,00	156,00	66,00	62,00	73,00	79,00	61,00	84,00	83,00
7	Media Canoa	Cauca	67,00	104,00	195,00	71,00	71,00	77,00	91,00	69,00	93,00	95,00
8	Rio Frio	Cauca	69,00	138,00	218,00	76,00	73,00	90,00	106,00	74,00	107,00	99,00
9	Guayabal	Cauca	74,00	140,00	231,00	84,00	74,00	94,00	111,00	79,00	102,00	101,00
10	La Victoria	Cauca	74,00	140,00	240,00	84,00	74,00	96,00	111,00	83,00	119,00	105,00
11	Ancaro	Cauca	75,00	140,00	245,00	89,00	75,00	101,00	118,00	88,00	125,00	111,00
12	La Luisa	Claro	0,70	1,20	2,00	0,80	1,40	1,50	1,60	0,80	2,00	0,40
13	Ortigal	Desbaratado	0,02	0,40	0,56	0,03	0,01	0,01	0,01	0,01	0,10	0,70
14	Buchitolo	Fraile	0,30	2,30	2,90	1,30	0,10	0,10	0,50	0,20	0,10	0,70
15	El Vergel	Guadalajara	0,70	2,30	2,30	1,20	1,30	0,40	1,60	1,00	1,50	0,10
16	Los Bueyes	Guengue	1,50	2,60	1,20	2,80	2,30	1,50	1,10	0,10	2,10	3,00
17	Potrerito	Jamundi	1,00	0,90	1,20	0,10	1,10	0,90	1,50	0,60	0,90	0,10
18	La Sorpresa	La Paila	0,54	2,58	1,98	0,73	0,50	0,64	0,78	0,47	0,84	0,16
19	Cabito	La Quebrada	0,30	0,10	0,10	0,20	0,20	0,10	0,40	0,10	0,30	0,30
20	Lomitas	La Teta	0,50	0,90	1,70	0,30	0,50	0,60	1,00	0,60	0,90	0,60
21	Caicedonia	La Vieja	15,00	15,00	28,00	11,00	12,00	12,00	16,00	11,00	15,00	12,00
22	Cartago	La Vieja	19,00	25,00	52,00	20,00	22,00	26,00	28,00	20,00	34,00	21,00
23	Media Canoa	Media Canoa	0,15	0,29	0,19	0,12	0,02	0,31	0,17	0,14	0,19	0,06
24	Santa Librada	Morales	0,30	2,00	1,60	0,10	0,10	0,10	0,40	0,20	0,60	0,64
25	Abajo	Ovejas	6,60	6,80	8,30	3,30	5,00	5,40	7,30	5,60	6,60	5,90
26	Pto. Tejada	Palo	6,90	6,00	18,00	5,70	4,30	5,90	8,60	5,60	8,10	9,60
27	Pichinde	Pichinde	0,40	0,50	1,00	0,20	0,70	0,60	0,60	0,60	0,80	0,70
28	Pte.Ferrocaril	Quinamayo	0,20	0,80	1,60	0,30	0,30	0,50	0,60	0,40	0,20	0,20
29	Timba	Timba	3,10	5,20	9,80	3,60	3,60	4,60	5,70	3,50	8,70	3,80
30	Meteguadua	Tulua	4,50	8,00	8,00	4,80	4,20	4,20	4,80	3,70	4,00	4,80

edures it was applied the number of turning points test. As pointed out by NERC (1975), the assumption of randomness can not be proved but may be disproved if features of non-randomness are found in the series. In a random sequence the total number of turning points is approximately Normally distributed (Yule and Kendall.,1950), with mean $\frac{2(N-2)}{3}$ and variance $\frac{(16N-29)}{90}$. Furthermore, it was carry out the serial correlation coefficient to check the effect of persistence in the series and its aim was to investigate the degree to which the discharge in one year is dependent upon the magnitude of the discharge in the preceding year. The lag one serial correlation has been taken as a measure persistence. It is defined as:

$$R_1 = \frac{\frac{1}{N-1}\left[\sum_{i=1}^{N-1}(X_i - \overline{X})(X_i - \overline{X})\right]}{\frac{1}{N}\left[\sum_{i=1}^{N}(X_i - \overline{X})\right]}$$

where R_1 is the serial correlation coefficient at lag one and X_i is the i^{th} value in the series. For a random sequence the value R_1 should be close to zero and it varies from zero only by sampling variation. According to Clarke (1973), R_1 is normally distributed with mean $\frac{-1}{(N-1)}$ and variance $\frac{(N-2)^2}{(N-1)^3}$. The above two statistical distribution free tests, namely, turning points and serial correlation coefficient implie that they can be used whatever the form of the distribution in the parent population may be . Such tests have the advantage that their approximate validity is comparatively easy to verify, and they generally do not require a lot of computation. They are often less powerful than other standard procedures but the loss of power is usually small and it is more than compensated for their wider applicability (Keeping.,1966).

Regional analysis

The hydrologist frequently needs to estimate event magnitude from sites which are either ungauged or have records of very short duration. When only a short record is available it is not advisable to choose a distribution at-site base on the sample alone (NERC (1975), Cunnane (1989)), but prior information about the form of the distribution e.g. any existing regional frequency curve must be used. Usually attempts are made to use the information at hand which at first sight is that belonging to neighbouring gauge catchments. The use of information belonging to a particular region to establish quantiles estimates at a site is called regional analysis. By using this procedure the data belonging to a particular region is assumed as being derived from the same population, so that it is comdined with the prior information into a single standardized sample.

The regional methods are performed in order to reduce the standard error of the estimate (Benson (1960), Dalrymple (1960)), to reduce the sampling error and, even for a gauge site, will produce more reliable event estimates than a single station frequency analysis, Kite (1977).

Accurate flood frequency analysis demands some form of regionalization, but to regionalize climatic variables and transform them to get estimates of the extreme quantiles of streamflow is a approach designed, presumably not by intent, to maximize the root mean square error of the final quantile estimates (Wallis 1980). Hence, the use of more than one set of data produces more robustness in the estimation of the parameters.

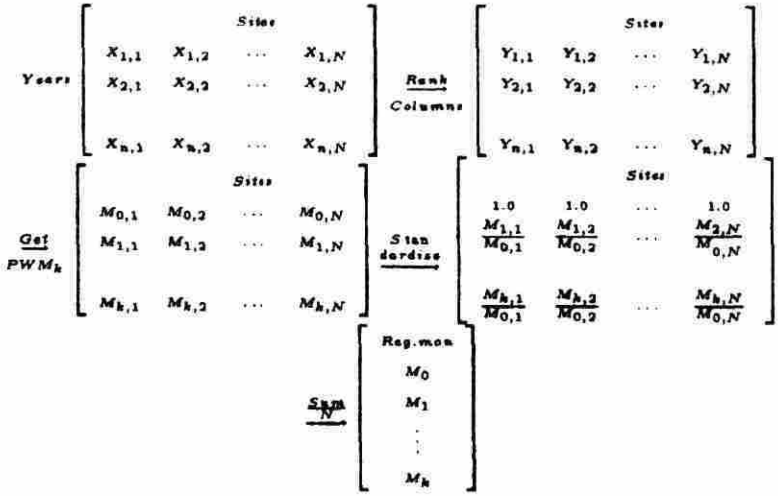
The most generalized scheme of regional analysis evolves the determination of a dimensionless flood or low flow frequency relationship, the estimation of a mean annual minimum or maximum flow series either from a record of data or from an equation linking it to catchment characteristics and the estimation of the quantile by using the relationship $\frac{Q}{Q} = Q_r$, where Q_r is the quantile variate, while Q is the estimate at site and $\overline{Q}$ is the mean annual minimum or maximum flow (González (1989)).

Different regionalization methods are available (Dalrymple 1960, NERC 1975, U.S. Water Resources Council 1977, Wallis 1980), The regional PWM was chosen in this study. Wallis (1980), indicated that this procedure may be found useful in those situations, where the records are short as this case. Besides, if the data in the region are indeed 'homogeneous 'or 'quasy-homogeneous ', then the above procedure can be expected to result in quantile estimates that are better than the comparable at-site estimate for all sites.

The regional PWM technique is particularly robust when the available record samples are either of very short length, highly kurtotic or skewed (Greenwood et al.,(1979), Greis and Wood (1981). Robustness studies done through Montecarlo simulation have shown the better performance of this methodology over other regional frequency algorithms (Lettenmaier and Potter (1985), Hosking et al., (1985a,b) and, Wallis and Wood(1985)).

Figure 2 show the procedure to obtain the original PWM as suggested by Wallis (1980).

Figure 2. **PWM regional Method as proposed by Wallis (1980)**



pfd Parameter Estimation

To estimate the parameters of the pdf's used, the regional PWM was applied. The PWM method was introduced by Greenwood et al.,(1979). It is useful for distributions which can be expressed in inverse form (X=X(F)), where F is the cumulative distribution function. It leads to simple and unbiassed estimators (Cunnane., 1985). The PWM's are defined as:

$$M_{l,j,k} = E[X^l F^j (1-F)^k] = \int_0^1 X^l F^j (1-F)^k dF$$

Where l, j, k are real numbers and $F = F(x) = P(X \leq x)$. It is observed that the above expression is a more general expression than conventional moments. If $j = k = 0$, $M_{1,0,0}$ represents the moment about the origin of order 1.

As it is possible to express the Gumbel, Wakeby, GEV, and Weibull distributions in inverse form, they were fitted by using the PWM method.

Fitting the distributions by PWM.

In the application of the PWM method of parameter estimation to the different distributions applied , the moment used varies. While for the GEV distribution , Hosking et al., (1985b) obtained the parameters by using the moment $M_{1,j,0}$,Greenwood et al., (1979) used the expression $M_{1,0,k}$, to estimate the parameters of the Weibull, Gumbel, Generalized lambda, Logistic, Wakeby and Kappa distributions.

Landwehr et al., (1979 b, c), found that PWM algorithm performs well for the Wakeby distributions using moderately biased estimate of $M_{1,0,k}$ given as:

$$M_{1,0,k} = \frac{1}{N}\sum_{i=1}^{N} X_i(1-P_i)^k$$

Where P_i denotes the i^{th} plotting position calculated as: $P_i = \frac{(i-0.35)}{N}$, i is the rank and N the sample size.

The same biased estimator was utilized by Landwehr et al ., (1979a) to estimate the Gumbel parameters. Similarly, Hosking et al., (1979a, b) also applied the same estimator in the determination of the GEV parameters. In order to obtain the Weibull parameters the same estimator was implemented in this study.

Details of the derivation procedure and parameter estimates in the term of PWM can be found in Greenwood et al., (1979), Landwehr et al., (1979a,b,c), Hosking et al., (1985a,b). The parameters of the regional distributions were obtained by regionally averaging PWM's, as outlined in Fig.2. Table 5 shows the cumulative distribution functions for the different distributions applied in this study and the values of the parameters for the distributions fitted in the region for low flows and AM flow series.

Table 5. **Parameter estimated for the pdf's used.**

Distribution	Cumulative Distribution Function CDF F(X) X(F)	Distribution Parameter Values									
		ε		α		λ		ζ		δ	
		Max	Min	Max	Min	Max	Min	Max	Min	Max	Min
Gumbel EV1	$\exp\left[-\exp\left\{-\left(\frac{x-\varepsilon}{\alpha}\right)\right\}\right]$	0.85	0.763	0.27	0.41	0	0				
general extreme value GEV	$\exp\left[-\left\{1-\lambda\left(\frac{x-\varepsilon}{\alpha}\right)\right\}^{\frac{1}{\lambda}}\right]$ $\lambda \neq 0$	0.66	0.73	0.36	0.33	0.86	-0.22				
Weibull (2)parameters	$1-\exp\left\{-\left(\frac{x}{\alpha}\right)^{\lambda}\right\}$	0	0	1.11	1.18	342	2.07				
Weibull (3)parameters	$1-\exp\left\{-\left(\frac{x-\varepsilon}{\alpha}\right)^{\lambda}\right\}$	0.347	0.398	0.758	0.635	3100	1.86				
Wakeby	$\varepsilon+\alpha\left[1-(1-F)^{\lambda}\right]$ $-\zeta\left[1-(1-F)^{-\delta}\right]$	0	0	0.739	0.802	11.073	31.96	14.601	4.47	0.02	0.193

Goodness of fit test's

Most of the studies which have attempted to discriminate between distributions in frequency analysis methods have relied at least to some extent on objective goodness of fit indices, i.e. Benson (1968), Bobee and Robitaille (1975), NERC (1975), U.S.W.R.C. and Beable and McKerchar (1982). However, it is acknowledged that classical goodness of fit indices such as chi square and Kolmogorov-Smirmov test are not sufficiently sensitive or powerful, and they are of little help in choosing any distribution, Benson (1968), Bobee and Robitaille (1975), NERC (1975). Wallis and

Wood (1985) pointed out that a goodness of fit test with short record can not provide unequivocal answers about the nature of the underlying distribution or necessary reliable guide to the problem of estimating extreme flow quantiles with a maximum accuracy, which emphasizes the need for robust estimation techniques.

Test's used

There were performed four goodness fit tests, namely, chi square, probability plot criteria (Benson 1968), 'DMAX ' and a graphical tests. In applying the chi square tests and the probability plot criteria the station year assumption in first made. The data of each sample is standardised by division by the sample mean. The values are pooled together into a single sample and it is assumed that this is a random sample from a single population.

Probability plot criteria (PPC)

The method consisted of evaluating how well each regional distribution under test fitted the whole of the standardized series. The following PPC index was chosen to estimate the deviations produced between the observed quantile (Q_i) and the variate value on the curve at the i^{th} plotting position.

$$PPC = \frac{1}{N}\sum_{i=1}^{N}|Q_i - E(Q_i)|$$

$E(Q_i)$ depends on both the parameters and form of the distribution under test and implies that the plotting position or its equivalent on the probability scale is used (NERC 1975). This involves some compromises because the exact values of $E(Q_i)$

for the GEV, Wakeby and Weibull distributions are unknown. However, distribution free Hazen formula $\frac{(i-0.5)}{N}$, is used here. The Gringorten formula $\frac{(i-0.44)}{(N+0.12)}$ was used for the EV1 distribution.

'DMAX ' goodness of fit test.

This tests developed by Njenga and Cunnane (1985), has a philosophical basis similar to the Kolmogorov-Smirmov test, but in this cases the advantages of the computer are used to make inferences from the regional distribution of the statistics (skewness coefficient Cs, coefficient of variation Cv, coefficient of kurtosis Ck).

The statistics generated from a distribution under test are compared with those of the observed values of annual minimum and annual maximum flow series. The region was simulated 100 times by the computer. From the above simulation was obtained the same number of statistics (Cv,Cs and Ck) for each station. In each repetition the statistics belonging to the stations in the simulated region were ranked from smallest to largest. The mean in every rank and for each statistics was obtained. The absolute derivations of the statistics from the average for each rank and repetition were computed. The maximum value for each repetition denoted as 'DMAX' was selected. Thus, it is obtained a 100 rows vector of 'DMAX 'values for each statistics. Figure 3 shows the procedure.

Figure 3. **algorithm to compute the 'DMAX' goodness of fit test**

$$\left[\ \textit{Distribution under test}\ \right] \rightarrow \left[\begin{array}{c}\textit{Generate samples of size } 10\\ \textit{for the region of } n \textit{ stations}\end{array}\right] \rightarrow$$

$$\left[\begin{array}{l}\textit{Calculate statistics } Cs, Cv, Ck\\ \textit{for each stations in the region}\end{array}\right] \rightarrow \begin{array}{c}\textit{rank the values for}\\ \textit{each repetition}\\ \textit{and statistics}\end{array} \xrightarrow[\textit{for } Cv]{\textit{sample}}$$

$$\left[\begin{array}{c}\underset{\longrightarrow}{\textit{sites } n}\\ \textit{repetitions} \downarrow\end{array}\right] \rightarrow \left[\begin{array}{cccc} C_{v_{1,1}} & C_{v_{1,2}} & \cdots & C_{v_{1,n}}\\ \cdots & \cdots & \cdots & \cdots\\ C_{v_{100,1}} & C_{v_{100,2}} & \cdots & C_{v_{100,n}}\end{array}\right] \rightarrow$$

$$\xrightarrow[\textit{columns}]{\textit{averaging}} \left[\begin{array}{c}\overline{C_{v_{1,1}}}\ \overline{C_{v_{1,1}}} \cdots \overline{C_{v_{1,1}}}\\ \textit{mean of ranked statistics}\end{array}\right] \xrightarrow[\textit{deviations}]{\textit{obtain absolute}}$$

$$\left[\begin{array}{ccc} |C_{v_{1,1}} - \overline{C_{v_1}}| & \cdots & |C_{v_{1,n}} - \overline{C_{v_n}}|\\ \cdots & \cdots & \cdots\\ |C_{v_{100,1}} - \overline{C_{v_1}}| & \cdots & |C_{v_{100,n}} - \overline{C_{v_n}}|\end{array}\right] \xrightarrow{\begin{array}{c}\textit{obtain Maximun deviation}\\ \textit{value } DMAX_i \textit{ for each}\\ \textit{repetition}\end{array}}$$

$$\left[\begin{array}{c} DMAX_1\\ \cdots\\ DMAX_{100}\end{array}\right] \xrightarrow{\begin{array}{c}\textit{fit the Maximun value}\\ \textit{to a GEV distribution}\end{array}} \begin{array}{c}\textit{Perform the test with the}\\ \textit{critical value obtain}\\ \textit{from the sample}\end{array}$$

The maximum deviation between the ranked statistic from the observed series and their corresponding means obtain in the simulation $\overline{C_{v_1}}$ *to* $\overline{C_{v_n}}$ is the goodness of fit index called 'DMAX'. If 'DMAX 'is greater than the critical value taken from the upper tail of the fitted distribution obtained with the DMAX simulated values, then the hypothesis that the series comes from the distribution under test is rejected. The significance level selected for this study was 5%.

Graphical test.

For each the distributions under test in both sets of series (annual maximum and minimum), 500 samples of size 10 were generated and each simulation was ranked in ascending order. Thus each of the order statistis from one to ten have 500 values.

Instead of assuming any distribution for the random variables belonging to each order statistic, they were individually ranked in ascending order. From every order statistics the 12^{th} and 488^{th} values were chosen as the lower and upper bounds corresponding to the middle 95% confidence interval for this test. The pair of points so chosen were plotted against their corresponding plotting position order statistics on ordinary graph paper. Two separate smooth curves were drawn through the points corresponding to the upper and lower bounds respectively.

The test consists of plotting for each value of order statistics between 1 and 10 the standardized ranked values for each station in the region. The visual inspection as well as the number of points outside the bounds helps to obtain the best approximation for the region.

Results

From the persistence and randomness test's applied to the annual minimum flow series, the station number 30 on Tuluá river was withdraw from the analysis because it displayed persistence and non-randomness. None of the AM flow series showed neither persistence nor non-randomness. As a result the analysis was carry out with 29 stations for analysis of annual minimum flow and 30 stations for AM flow.

Before any goodness of fit test was carry out the Weibull three parameters distribution fitted to the annual minimum flow series was dropped from the analysis because its lower bound ($\mu = 0.395$) (see Table 5) was greater than 10% of the 290 standardized annual minimum flow values. In regard to the annual maximum flow series, it was observed that the value of the GEV distribution shape parameter (λ=0.061) is small and positive which corresponds to a EV3 distribution. Hosking et al., (1985a), developed a test to see whether the shape parameter λ is zero or not. The test was carry out and did not reject the hypothesis of $\lambda = 0$ at a significance level of 5%. As a consequence, the GEV distribution was dropped from the analysis, and the EV1 distribution was the only distribution from the GEV family remaining.

Table 6. PPC **Index applied to the probability plot.**

Annual series	Distribution	PPC Index
Minimum	EV1	0.21
	GEV	0.24
	Wakeby	0.24
	Weibull	0.42
Maximum	EV1	0.16
	Wakeby	1.01
	Weibull(2 parameters)	0.36
	Weibull(3 parameters)	0.29

Annual minimum flows

The chi-square goodness of fit test shows the GEV distribution as the only distribution accepted at 5% level which describes the annual minimum flow series. In regard to the chi square index it is observed that the EV1 and Wakeby distribution have a similar index, while the Weibull distribution display a greater value. The hypothesis that the sample comes from the above distributions was rejected. This result somehow unexpected since the Wakeby distribution ought to be sufficiently flexible to fit any sample well.

Using the 'DMAX 'goodness of fit test. The regional GEV distribution perform better than the other fitted distributions. The Wakeby and EV1 distributions behave similarly in this test. The worst performance was obtained with the Weibull distribution. The hypothesis that the statistics Cv,Cs and Ck comes from the GEV distribution is accepted, while in the other distributions the hypothesis that the Cv comes from the distribution under test is rejected.

Fig.4 and Table 6 shows the data with the distribution under test displayed in EV1 paper plots. From a visual inspection of the plots and especially the lower tail of the distributions, it is observed that the fit in this part of the distribution, the most important for low flows, is not quite acceptable in any of them. The order of best fit according to the PPC criteria shown is : 1.)EV1, 2.)GEV, 3.)Wakeby and 4.)Weibull distribution.

The performance of the distributions (see Fig. 5 and Table 7) when is used the graphical test has the following order : 1.)EV1,2.)Wakeby, 3.)GEV and 4.)Weibull distribution.

From the above goodness of fit, it is observed that GEV distribution, followed by the EV1 distribution performed in the most stable way throughout all the tests.

Lettenmaier (1985), pointed out that two parameter distributions, as in this case the EV1 distribution, when they are used with an appropriate regionalization method, can result in quantile estimates with quite low variability but at expenses of considerable bias, and the GEV/PWM appear to yield quantile with low variability and low bias. That is why it is chosen the regional algorithm GEV/PWM for annual minimum flow in the region.

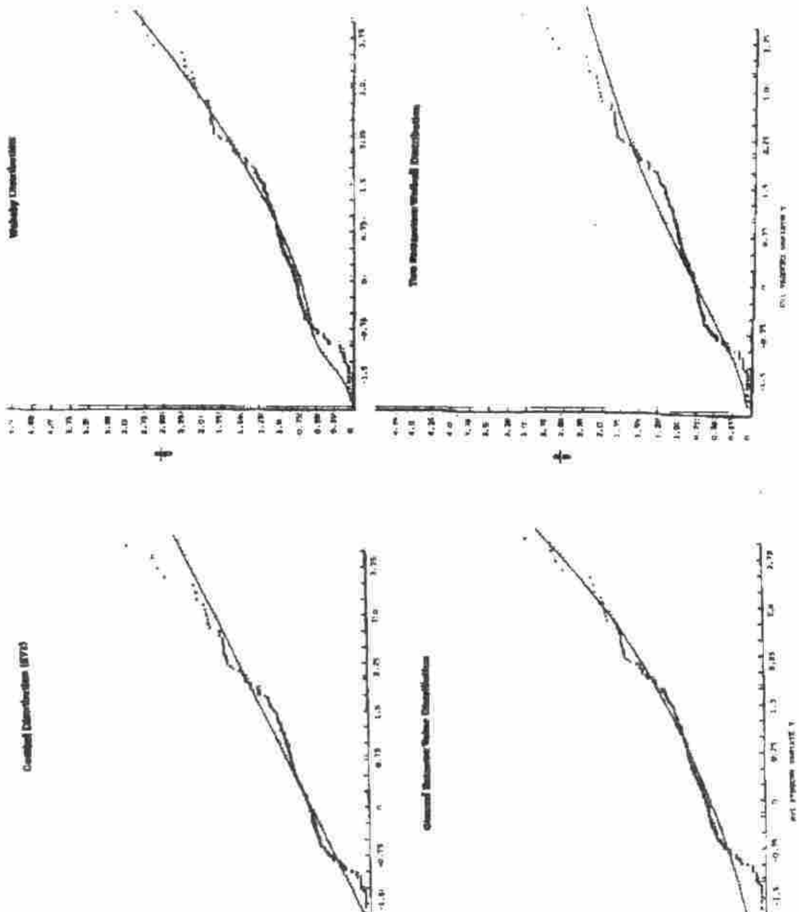


Fig.4. Probability Plot test for the Annual Minimun Flow series

Fig.5. Graphical test for the annual minimun flow series

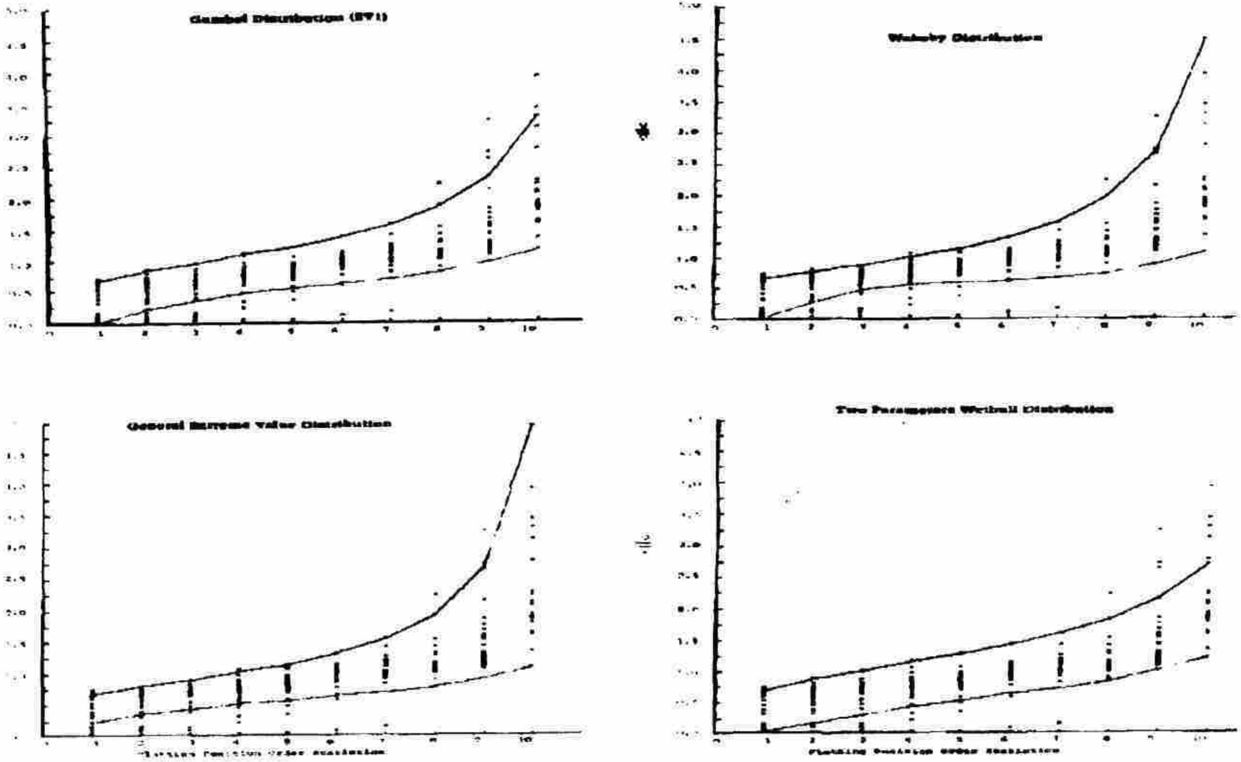


Table 7. **Number of values belonging to the series outside of the probability range.**

Annual series	Distribution	Number of values
Minimum	EV1	23
	GEV	29
	Wakeby	28
	Weibull	31
Maximum	EV1	7
	Wakeby	9
	Weibull(2 parameters)	13
	Weibull(3 parameters)	9

Annual maximum flows

The result for the chi square test shows that EV1 distribution is the only one for which the hypothesis that the AM series comes from the distribution under test is accepted, the significance level used was 5%.

The performances of EV1 and Wakeby distribution in the DMAX test are similar. It is accepted the hypothesis that the statistics Cv,Cs and Ck came from this distributions. The Weibull three parameter distribution followed by the Weibull two parameter distribution obtained the worst performance.

The probability plot test showed that performances of the EV1 and Wakeby distributions are similar in the upper tail of the plot. However, the PPC criteria indicated a better performance of EV1 distribution, followed by the Weibull three parameter distribution and the Weibull two parameter (see Fig.6 and Table 6).

Fig.7 together with Table 7, showed again that the EV1 and Wakeby distribution perform in a similar way in the graphical test followed in the same order for the Weibull three and two parameter distributions.

From the results, it is observed that when it is applied to annual maximum flow series the EV1 distribution performed better that any other distributions. Therefore, the EV1 distributions is chosen for the annual maximum flow series.

NERC (1975) pointed out that it has not been shown conclusively that the two parameter distribution are inapplicable, moreover the performance of this algorithm EV1/PWM has shown that quatiles produce low variability despite doubts about the homogeneity of the region. Greis and Wood (1981).

Conclusions

The main drawback in the choice of the distributions was that none of the test. except the probability plot, put emphasis in the lower or upper tail of the distribution which are the part of interest in low or high flow studies.

As was stated by Wallis (1980), theoretical innovations have not value for their own sake. They are only valuable if it can be proven that the new technique is in some sense superior to other techniques already available and accepted. This results may have value if it is compared with other techniques applied in Colombia and helps towards a procedure of regional estimates based upon the most update regional statistical techniques such as the cluster analysis.

Finally, the distributions chosen were the GEV distribution and EV1 distribution for annual minimum flow and annual maximum flow series respectively.

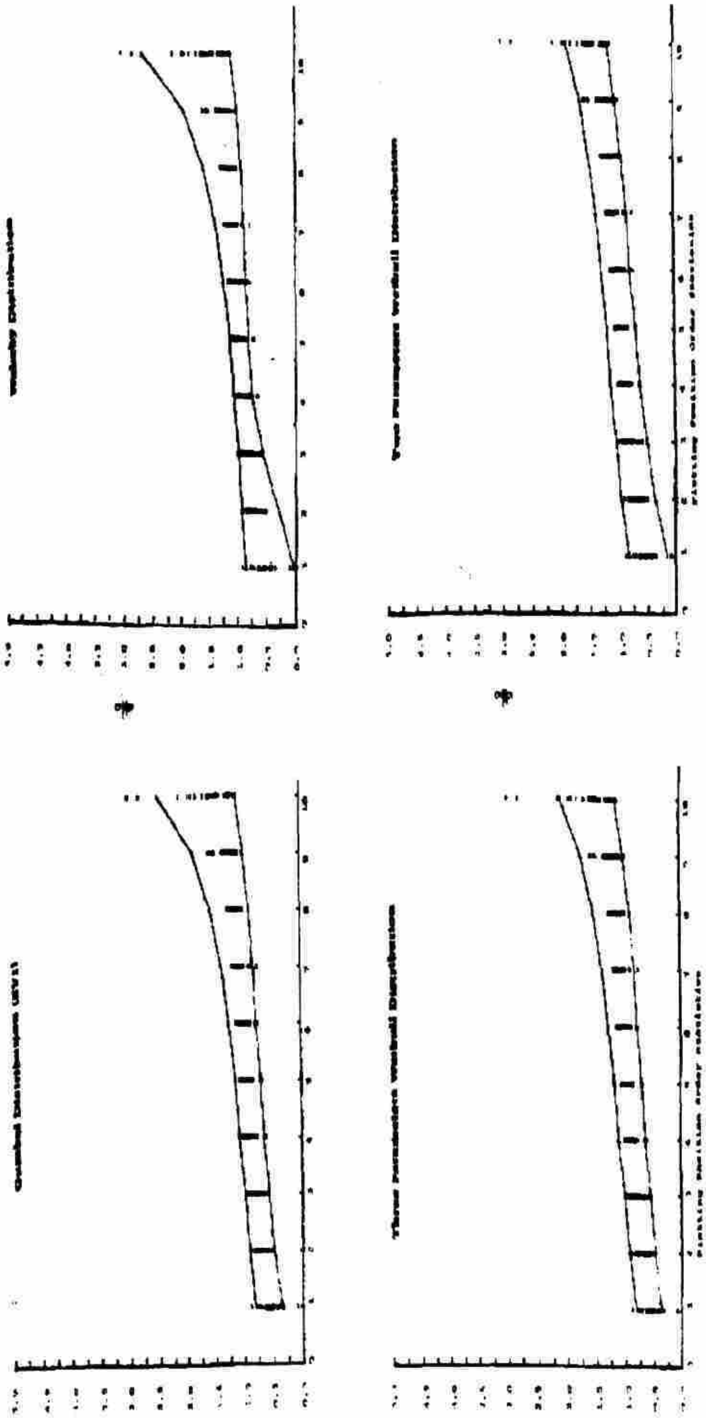


Fig. 7. Graphical test for the annual maximun flow series

King, M., (1985). *A low flow frecuency analysis for Irish river.* M. Sc. thesis, University College Galway, Ireland.

Kite, G.W.,(1977). *Frequency and risk analysis in hydrology.* Water Resources. Publications, Fort Collins, Colorado. U.S.A.

Lanwehr, J.M., Matalas, N.C. and Wallis, J.R. (1979(a)). *Probability weighted moments compared with some traditional techniques in estimationing Gumbel parameters and quantiles.* Water Resources Research, Vol 15 (5), 1055-1064.
Landwehr, J.M, Matalas, N.C. and Wallis, J.R. (1979 (b-c)). *Estimation of parameters and quantiles of Wakeby distribution.* Water Resources Reseach, Vol 15 (6), 1361-1379.

Lettenmaier, D.P., (1985). *Regionalization in flood frequency analysis: -Is it the answer?*.Paper pres. at U.S.-China Bilateral symposium on the analysis of extraordinary flood events, Nanjng.

Lettenmaier, D.P., and Potter, K.V., (1985). *Testing flood frecuency estimation methods using a regional food generating model.* Water Resources Research, Vol.21(12), 1903-1914.

NERC, (1975). *Flood studies report.* National Environment Resources Council, Vol 1-5, 1100 pp.

Njenga, M., (1985). *Simulation applied to the inference problem of the underlying distribution of hidrologic random variables* . M.Sc. Thesis, University College Galway, Ireland.

U.S.W.R.C. (1977). *Guidelines for determining flood flow frequency.* Bulletin No 17A of the Hidrology Committee. Washington.

Wallis, J.R., (1980). *Risk and uncertainties in the evaluation of flood events for the desing of hydrologic structures.* Keynote address at "Seminar on Extreme Hydrological Events-Foods and Droughts "Erice, Italy, 33 pp.

Wallis, J.R. and Wood, E.F. (1985). *Relative accuracy of log-Pearson III procedures.*A.S.C.E. Journal of Hydraulic Engineering, 111 (7),1043-1056 (See also discussion and reply, 113(9), 1205-1214, 1987).