Genetic Programming Applied for Shaping a Design Hydrograph from the Historical Flows' Pattern

Dr. **Maritza Liliana ARGANIS JUÁREZ[1,2],** M. A. **Juan José BAÑOS MARTÍNEZ[3]**, M. I. **Margarita PRECIADO JIMÉNEZ[4],** Ing. **Cecilia GONZÁLEZ CORREA[1]**

[1] Instituto de Ingeniería, Universidad Nacional Autónoma de México; MArganisJ@iingen.unam.mx

[2] Facultad de Ingeniería, Universidad Nacional Autónoma de México; MArganisJ@iingen.unam.mx, cgonzalezco@iingen.unam.mx

[3] Facultad de contaduría y administración, Universidad Nacional Autónoma de México; preciado@tlaloc.imta.mx

[4] Instituto Mexicano de Tecnología del Agua; juanjosebm@yahoo.com.mx

Abstract: *A method based on applying genetic programming (GP) in the ascending and the descending branch of a parameterized hydrograph is proposed in order to get a design hydrograph. Parameterization is done by considering peak flow, the peak time and the approximate time base estimated from the behavior of the annual maximum historical floods envelope and historical annual maximum floods envelope measured in the study site in the study site. With GP, was obtained the hydrograph's behavior before and after the peak flow for the historical data. Additionally, was made comparison of the GP results with a polynomial interpolation of Lagrange. Afterwards a design hydrograph was obtained, for "El Infiernillo" dam, considering the peak design flow, calculated with the Instituto de Ingeniería (IINGEN) Method. The volumes approached with GP was different from that reported by the IINGEN Method, but this difference could be corrected applying a factor to the hydrograph ordinates, with exception of the peak flow to keep the volume. Hydrograph shape obtained using GP and the IINGEN method were similar, with smoothed shapes in hydrograph obtained with GP, but ensuring similar shapes than the historical floods, even the volume, peak and base time.*

Keywords: *Average daily inflows, Base time, Design flood, "El Infiernillo" Dam, Genetic programming, Lagrange polynomial interpolation*

1. Introduction

There are different procedures to obtain the peak flow design, depending on the basin size, the measured data (if it comes to rainfall or runoff), and the basin physiographic information. The determining the behavior of design flow throughout the flood duration (time base) has been the subject of study by different authors: Pegram and Deacon (Pegram & Deacon 1992) worked by Hiemstra and Francis, they using the Pearson hydrographs type (Hiemstra & Francis 1979; Jiménez 2000). IINGEN Method based on alternating blocks and an analysis of annual maximum mean daily flow for a duration equal to the time base of the desired design flood (Domínguez et al. 2012) order spectral to get the shape of the hydrograph (Fuentes et al. 2015) historical increase envelope normalized method using a factor (Arganis et al. 2013), Hermitian hydrographs (Ramírez et al. 2000; Domínguez et al. 2012) among others. Many of these procedures are based in the case of floods that have a one peak, but sometimes provide different forms from those historically measured.
Moreover, the genetic programming (GP) is a random algorithm and it has tools of evolutionary computation for obtaining mathematical models from data of a dependent variable and n independent variables available and there have been several applications of it in problems of hydraulic and hydrological engineering, particularly in the rain runoff processes (Savic et al. 1999; Whigham et al. 2001; Goldberg & Holland 1988; Dorado et al. 2003; Rabuñal et al. 2007; Nourani et al. 2012; Danandeh et al. 2013; Meshgi et al. 2015).
Another deterministic algorithm that is useful to adjust tabular data functions is the Lagrange polynomial interpolation (Luthe et al, 1994; Chapra 2000) and it is applied in this work too to make some comparisons between results.

In this journal genetic programming (GP) was applied to parameterized data of annual maximum historical floods envelope of "El Infiernillo" Dam; located at Michoacán State in México; parameterization was performed on the values of time and flow taking into account the peak time (assumed in the center of all historical floods), the time base (was 15 days of duration) and peak flow of the historic floods (simultaneity assumed by matching the peaks in each year). Applying genetic programming with the simplest operators of addition, subtraction and multiplication polynomial forms were obtained for the ascending and descending branches of historical parameterized envelope. The factors obtained for each array are multiplied by the design peak flow and thus the flood is constructed. The procedure was repeated generating Lagrange polynomials. The results were compared with the flood previously calculated with the IINGEN Method. Similar hydrographs were obtained with all methods; only the volume obtained with the method GP was slightly higher than the IINGEN method, this was achieved by estimating an adjustment fix factor applied to the hydrograph ordinates. The procedure is simple and can be applied in different basins for validation purposes. This procedure is relevant because there are few international studies related to give the design hydrograph shape by considering both variables: volume and peak flow. Having this information another important hydraulic studies such as two dimension flow simulations in order to know flood plain can be performed.

So many countries, even Mexico, don't count with references that stand out procedures to give the estimation of the design hydrograph shape inasmuch as they only make simplification to regular geometric shapes (triangles, for example).

2. Methods

2.1 Study area and data descriptions

"El Infiernillo" dam was built from 1960 on the Balsas River in Michoacán State, Mexico; is part of the dams system allocated along this river together with "La Villita" and "El Caracol" see Figure 1. The dam has a surface reservoir of 108 000 000 m2 and its spillway was designed for an outflow of 38 800 m3/s. Hydroelectric power generation began in October 1964, during 1965 were installed four units of the first stage and by mid-1975 came into operation the two turbines of the second stage. At the end of September 1967 occurred an extraordinary growing in the Balsas River, with entrance flow of 25 200 m3/s due to Hurricane Beulah and total volume of 7 500 million of m3, which was regularized at a maximum outflow of 7 500 m3/s; it was necessary to operate partially open radial gates to reduce the discharge of pouring the indicated value in order to protect another localized dam downstream (Jose Ma. Morelos) that was in construction (Domínguez et al. 2014).

The sum of four hydrometric stations: Los Pinzanes, Panches, La Pastoria and Caimanera were considered as input date. For the years 1965 to 1994, these data were multiplied by a factor of 1.3, such factor was obtained with the comparison between the sum and data reported in the common period for "El Infiernillo" dam. For the period of years 1998 to 2013 the daily average flow of total inflows reported by the "Comisión Federal de Electricidad (CFE) for the operation dam were used (Domínguez et al. 2014; Gómez 2015).

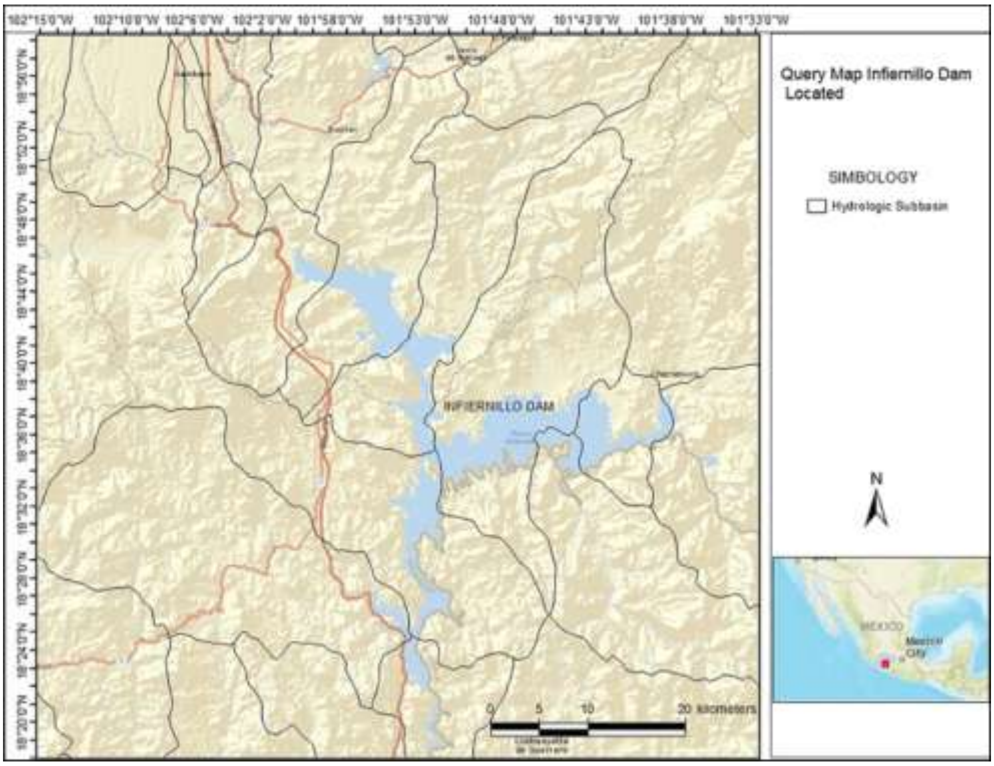

Fig. 1. "El Infiernillo" dam Michoacán, México

The Table 1 represent the daily average flow of total inflows reported as annual maximum flow, this comes from measured average daily inflows of the analyzed reservoir (records of Mexico's hydrometric stations are obtained from, first, the National Databank Surface Water (BANDAS) of the Water National Commission (CONAGUA 2016) and, second, the daily average flow of total inflows reported by the "Comisión Federal de Electricidad" (CFE).

Table 1: Annual maximum flood data. (CONAGUA 2016) and (CFE)

t, year	Q, cubic metre per second	t, year	Q, cubic metre per second
1965	4254.61	1991	5928.7
1965	4254.61	1991	5928.7
1966	2550.66	1992	2713.76
1967	14109.1	1993	2152.94
1968	2681.26	1994	2550.06
1969	5940.1	1995	5069.59
1970	3671.7	1996	3674.3
1971	5603	1997	1754.57
1972	2905.35	1998	5513.89
1973	7142.7	1999	3019.68
1974	2897.5	2000	3572.92
1975	4740.2	2001	4063.66
1976	9720.6	2002	3510.42
1977	3162.5	2003	5949.07
1978	3230	2004	8118.06
1979	2264.75	2005	3740.74

t, year	Q, cubic metre per second	t, year	Q, cubic metre per second
1980	2591.29	2006	3147.41
1981	4203.2	2007	2827.88
1982	1088.11	2008	3519.05
1983	2059.3	2009	2106.23
1984	7408.7	2010	3737.44
1988	2344.28	2011	5242.71
1989	1816.77	2012	2282.33
1990	2688.33	2013	15207.138

2.2 Study Annual maximum floods parameterized envelope curve

To construct the annual maximum historical floods envelope is required the measured average daily inflows of the analyzed reservoir (for Mexico's hydrometric stations, these records are obtained from the National Databank Surface Water (BANDAS) of the Water National Commission (CONAGUA) (CONAGUA 2016), with the hydrometric key and knowing the study area (hydrological region) see the Figure 2 and the Table 2, these data should be complemented with the information available of reservoir operation.

Fig. 2. "El Infiernillo" BANDAS interface (CONAGUA) query maps (CONAGUA 206)

Table 2: Annual maximum flood data. (CONAGUA 2016) and (CFE)

Code	Name	Flow	Basin	State
18481	La Caimanera	Balsas river	Balsas river	Michoacan
18487	Los Pinzones	Tacambaro river	Balsas river	Michoacan
18494	Los Panches	Tepalcatepec canal	Balsas river	Michoacan
18495	La Pastoria	El Marques river	Tepalcatepec river	Michoacan

After reviewing the data reported, the maximum annual flow is selected for each year and a certain number of days before and after the maximum flow, is set to define the time base and the maximum annual flood hydrograph's ascending and descending branches.
To have the worst behavior the peak flow of each annual flood are set matching them in peak time. In this study a time base of 15 days was considered for the historical floods and peak flow was placed on day eight.
For each day the maximum ordinate of all years is obtained, getting the annual maximum historical floods envelope.

The parametres are set for the time in the ascending branch considering the relation of time by the peak time (t/t_p) and the relationship of flow and peak flow (Q/Q_p); for the descending branch behavior was considered like the Hermitian hydrographs way, that is with $(t-t_p)/(t_p-t_b)$ and (Q/Q_p).

2.3 Genetic programming (GP)

Genetic programming (GP) (Koza 1992) takes place in a few years following the emergence of genetic algorithms (Goldberg 1989) in order to build computer programs and mathematical models with evolutionary random algorithms used as optimization methods. The genetic programming algorithm includes the establishment of the independent variables and the dependent variable in the problem, operators and constant vector to be considered for the construction of the models to be tested must also be defined. It should provide a probability of exchange or cross (crossing) of the best individuals (set of selected operations) and a probability of mutation must be given. A number of generations (iterations) is proposed to finish the optimization process. An example of the cross operator between two individuals is presented in Figure 3 were part of their nodes are exchanged each other, resulting in two new individuals.

In this study objective function consisted in minimizing the mean square error between the measured (Q/Q_p) data and the calculated with the models tested by the GP algorithm.

The problem starts with the random generation of an initial population of *n* individuals (each individual corresponds to a mathematical model consisting of different operators, variables and constants), individuals are then evaluated in the objective function and the best ones are selected (selection can be performed by obtaining a relative frequency of the result obtained by each individual in the objective function divided by the average value given by all tested individuals), individuals with higher relative frequency can be used more than once to be exchanged or crossover, and mutation may also create new individuals and the individuals with lower performance are eliminated and no enter to the exchange process and or mutation; so that the new population is again size *n*. The new individuals are again tested on their performance, selected and the best creates new individuals who pass to the next generation, this process is repeated until the number of generations or iterations is reached and the best individual in the last generation, will be the one with higher performance and represents the optimal mathematical model found.

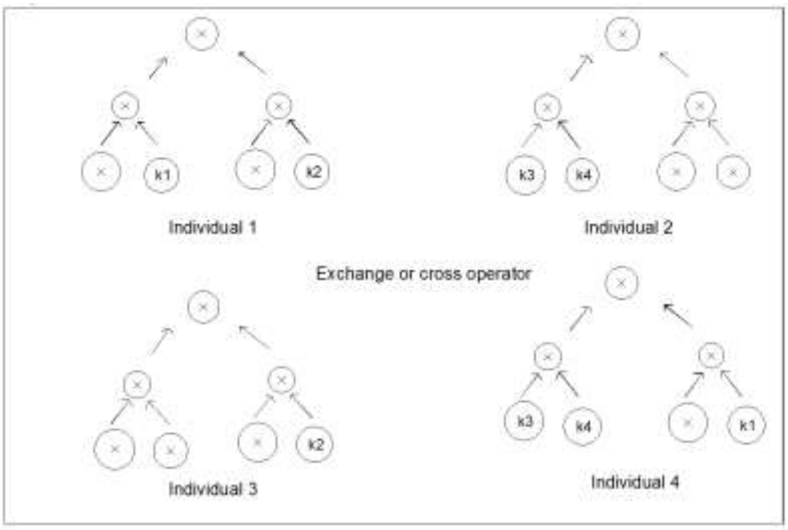

Fig. 3. Example exchange or cross (crossing) operator in GP

In this study the set of arithmetic operators *TS = [+,-,*]* was considered, a vector of constants obtained randomly, independent parameterized variable (t/t_p) in the upward branch and $t' = (t-t_p)/(t_b-t_p)$ in the descending branch and a parameterized dependent variable (Q/Q_p) was considered too. Populations of 300 individuals (models) of 25 nodes (consisting of operators, variables and

constants), a crosses a probability of zero point nine and mutation probability of zero point five, were considered; finally, 10 000 generations to complete the process were considered.
The GP programming used in this study was developed at the "Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas" (IIMAS, UNAM).

2.4 Lagrange polynomial interpolation

Lagrange interpolation allows a polynomial of degree *n* passing through *n + 1* points; the polynomial is of the form (Luthe *et al*. 1994), as show in formula (1) .
Parentheses on the far right margin of the page, as formula (1):

$$P(x) = a_1 x^N + a_2 x^{N-1} + ... + a_N x + a_{N+1} \quad (1)$$

the coefficients $a_1, a_2, ..., a_N$ are obtained by ensuring that the polynomial passes through each of the points $(x_1, y_1), (x_2, y_2), ..., (x_{N+1}, y_{N+1})$. It is a simple method to be applied, but its disadvantage is that to interpolate or extrapolate the result data corresponds to the image of the polynomial and may not correspond with the expected behavior of the analyzed function.

In this work it was used the code in Matlab polynomial interpolation of Lagrange presented in the website "Renewable energy of the University School of Industrial Engineering of Eibar (Energías renovables de la Escuela Universitaria de Ingeniería Técnica Industrial de Eibar)" (University of the Vasco Country 2016).

2.5 Mean square error and nonlinear determination coefficient

The mean square error (*MSE*), it is given by formula (2) and the nonlinear determination coefficient (R^2) for the formula (3), were applied in order to verify the goodness of the GP and Lagrange results.

$$MSE = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{Q}{Q_p} - \frac{\overline{Q}}{\overline{Q}_p}\right)^2 \quad (2)$$

Where *n* is the number of data, $\frac{Q}{Q_p}$ and $\frac{\overline{Q}}{\overline{Q}_p}$ are, respectively, the measured and calculated parameterized flow data, and with the formula (3) is calculated the nonlinear determination and correlation coefficient

$$R^2 = \frac{\operatorname{var}\left(e^{\frac{Q}{Q_p}}\right) - \operatorname{var}\left(e^{\frac{\overline{Q}}{\overline{Q}_p}}\right)}{\operatorname{var}\left(e^{\frac{Q}{Q_p}}\right)} \quad (3)$$

Where R^2 is the non linear determination and correlation coefficient; *var* is the variance operator; Q/Q_p is de parameterized flow and $e^{Q/Qp}$ is the error in the parameterized flow obtained with the tested model .

3. Results and discussion

From the historical record of maximum annual floods a time base of 15 days was selected. The annual maximum historical floods envelope was obtained by setting the peak flow for each flood in the same peak and the maximum of all flow was calculated for each day, see the first column of the Table 3. The parameterized time and flow of the envelope are shown in Table 3, columns three, four and five.

Table 3: Annual Parameterized envelope annual maximum flood

t, days	Q, metre cubic per second	t / t_p, dimensionless	$(t-t_p) / (t_b-t_p)$, dimensionless	Q / Q_p, dimensionless
1	2808.44	0.125		0.1847
2	3857.35	0.250		0.2537
3	3738.70	0.375		0.2459
4	3861.90	0.500		0.2540
5	6191.80	0.625		0.4072
6	6939.20	0.750		0.4563
7	6984.70	0.875		0.4593
8	15207.14	1.000		1.0000
9	14117.44		0.1429	0.9283
10	11470.92		0.2857	0.7543
11	5920.39		0.4286	0.3893
12	4382.91		0.5714	0.2882
13	3994.80		0.7143	0.2627
14	3942.83		0.8571	0.2593
15	4163.04		1.0000	0.2738

3.1 Annual maximum floods parameterized envelope curve with genetic programming

Ascending branch equation

Parameterized data envelope ascending branch were selected and genetic programming algorithm was applied giving the polynomial model, as formula (4):

$$\frac{Q}{Q_p} = 0.3014794 * \left(t/t_p \right)^3 + 0.166641861 * \left(t/t_p \right)^2 + 0.24158976_{N+1} \quad (4)$$

the coefficients $a_1, a_2, ..., a_N$ are obtained by ensuring

Descending branch equation

Moreover, descending branch data were considered and Equation (5) was obtained with genetic programming, as formula (5):

$$\frac{Q}{Q_p} = t'^3 - 0.4786217 * t'^2 - 1.287393114 * t' + 1.057715508 \quad (5)$$

the formula (5) is applicable in interval $t_p<t<t_b$; where, $t' = (t-t_p)/(t_b-t_p)$.

For t_p, the peak flow is ensured with the formula (6):

$$\frac{Q}{Q_p} = 1 \quad (6)$$

the formula (6) is valid at t = t_p.

Lagrange polynomials

Ascending branch equation peak flow for ascending branch data were selected up (eight points) and it was used a program implemented in Matlab to obtain the coefficients of the polynomial of Lagrange (grade seven) for the ascending branch, see the formula (7):

$$\frac{Q}{Q_p} = -652.0*\left(t/t_p\right)^7 + 2682.70*\left(t/t_p\right)^6 - 445.40\left(t/t_p\right)^5 + 3810.10*\left(t/t_p\right)^4 - -1803.00*\left(t/t_p\right)^3 + 464.90*\left(t/t_p\right)^2 - 59.40*\left(t/t_p\right) + 3.10 \quad (7)$$

the formula (5) is applicable in interval $t_1<t<t_p$;

Descending branch equation

Subsequently descending branch data (seven points) and the Matlab program was applied to obtain the Lagrange polynomial coefficients (grade six) for the descending branch, see the formula (8):

$$\frac{Q}{Q_p} = -141.1084t'^6 + 532.2911*t'^5 - 801.8006*t'^4 + 609.3198t'^3 - -239.849t'^2 + 43.249t' - 1.8282 \quad (8)$$

the formula (8) is valid in interval $t_p<t<t_b$; where, $t' = (t-t_p)/(t_b-t_p)$.

$$\frac{Q}{Q_p} = 1 \quad (9)$$

formula (9) applies at $t = t_p$

Substituting the parameterized time in each case the factors in Table 4 are obtained.

Table 4: Factors (Q/Q_p) GP and Lagrange for the desing hydrograph

t, days	GP	Lagrange
1	0.2448	0.2220
2	0.2567	0.2915
3	0.2809	0.2821
4	0.3209	0.2859
5	0.3803	0.4322
6	0.4625	0.4723
7	0.5711	0.4657
8	1.0000	1.0000
9	0.8669	0.9283
10	0.6741	0.7543
11	0.4968	0.3893
12	0.3524	0.2882
13	0.2584	0.2627
14	0.2323	0.2593
15	0.2917	0.2738

3.2 Mean square error and nonlinear determination coefficient

In the Table 5 are set the mean square errors obtained with GP and Lagrange methods for the ascending and descending branch (AB and DB).

Table 5: Mean square error obtained in ascending and descending branches (AB and BD)

GP AB	GP DB	Lagrange AB	Lagrange DB
0.0028	0.0038	0.0008	0.0038

In Figure 4, the measured and calculated points of the ascending and descending branch (AB and DB) were drawn against the perfect fit and the linear R^2 obtained in an Excel© worksheet is given too. Tables 6 and 7 show the variance of parameterized Q/Q_p and the error obtained with both GP and Lagrange models in the ascending and descending branches. Finally in the Table 8 the nonlinear determination and correlation coefficients are shown.

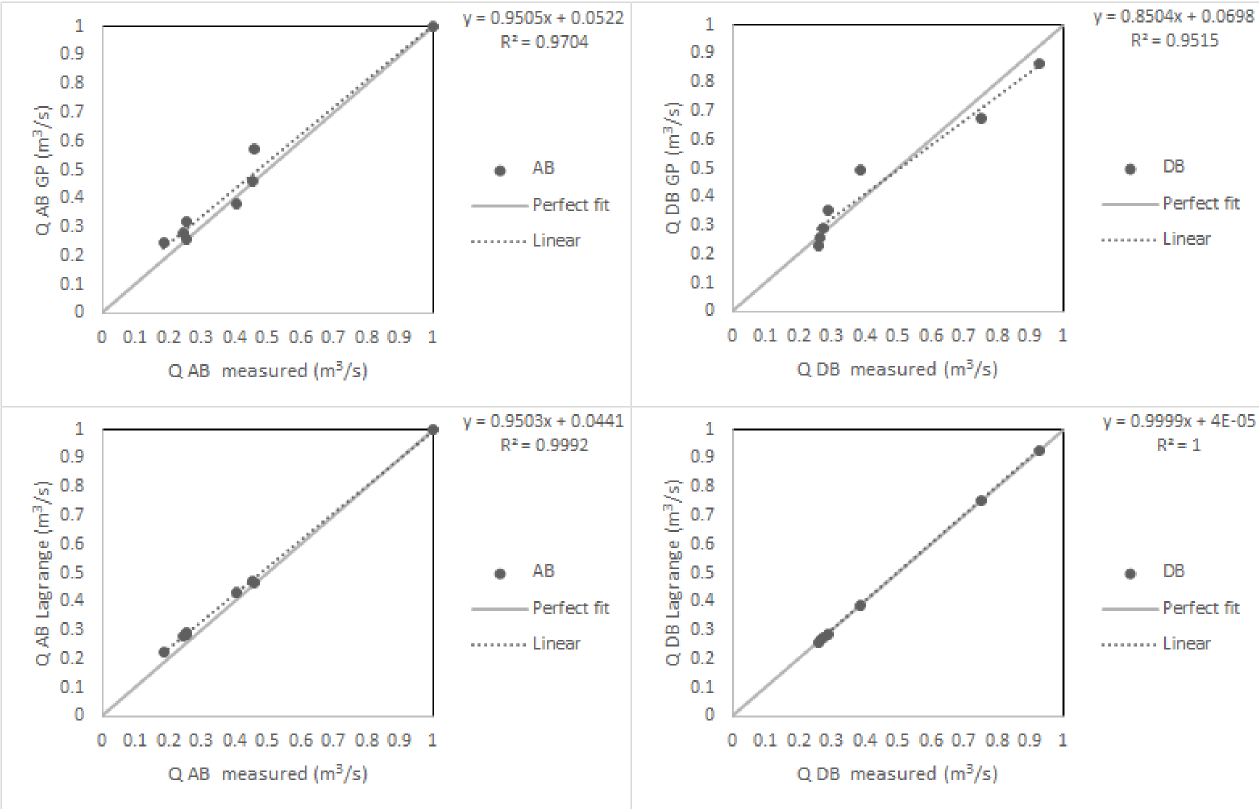

Fig. 4. Measured and calculated data with GP and Lagrange models in ascending and descending branches (AB and DB, respectively) against the perfect fit.

Table 6: Variances for ascending branch (AB)

Parameter	GP AB / Var Q	Var e_Q	Lagrange AB/ Var e_Q
	0.0684	0.0021	0.0002
R^2 nonlinear	0.9700		0.9968
r	0.9849		0.9984

Table 7: Variances for descending branch (DB)

Parameter	GP DB /Var Q	Var e_Q	Lagrange DB/ Var e_Q
	0.0756	0.0045	0.0045
R^2 nonlinear	0.9407		0.9408
r	0.9699		0.9699

Table 8: Nonlinear determination and correlation coefficient (R^2 and r)

Parametre	GP AB	GP DB	Lagrange AB	Lagrange DB
R^2	0.9700	0.9407	0.9968	0.9408
r	0.9849	0.9699	0.9984	0.9699

By multiplying the peak flow Q_p by design factors the hydrographs with the two methods are obtained and this are shown in the column 2 and 3 in the Table 3. The obtained floods were compared with both IINGEN Method and the increase of the historical maximum flood envelope, see the Table 9 and the Figure 4.

Table 9: Comparison of design hydrographs the annual maximum historical floods envelope

t, days	GP Q, m^3/s	Lagrange Q, m^3/s	Max envelope Q, m^3/s	IINGEN Q, m^3/s
1	9116.92	8269.82	6878.38	4990.87
2	9561.37	10857.98	9447.35	10140.71
3	10462.95	10505.14	9156.75	8589.49
4	11953.24	10649.75	9458.49	8616.84
5	14163.82	16098.97	15164.83	9005.45
6	17226.28	17590.45	16995.35	17885.25
7	21272.22	17344.29	17106.79	31500
8	37245.02	37245.02	37245.02	37245.02
9	32289.57	34575.85	34576.16	34754.98
10	25108.41	28094.25	28094.35	25700
11	18502.68	14500.24	14500.10	10636.63
12	13123.90	10734.99	10734.54	5701.63
13	9623.58	9784.81	9783.98	9379.9
14	8653.24	9657.94	9656.69	6456.25
15	10864.40	10197.69	10196.02	6776.73
16	21528.08	21263.66	20649.15	19645.61

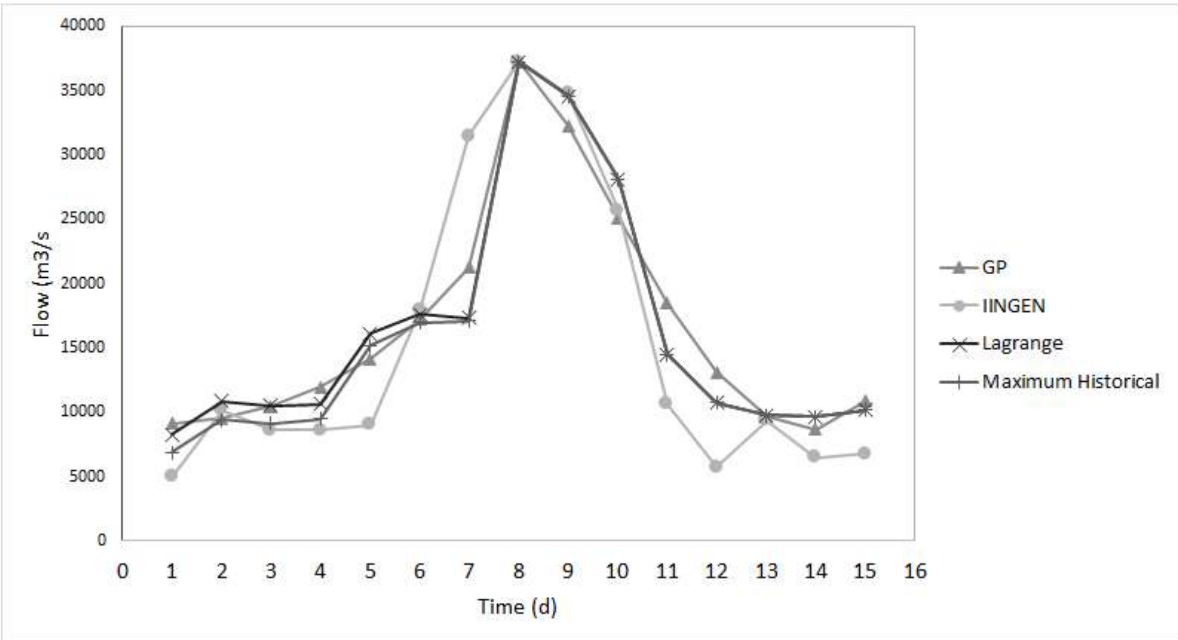

Fig. 5. Comparison of design hydrographs. Tr = 10 000 years. "El Infiernillo" dam.

In the Table 3 and Figure 5 shows that the method GP reports a volume slightly greater than those obtained with Lagrange polynomials respect to the design hydrograph calculated with the historical maximums envelope whereas the method of the Instituto de Ingeniería reported a lower volume.

Using the Solver tool © Excel © and the generalized reduced gradient method (GRG nonlinear), a factor to affect the calculated GP hydrograph was obtained to preserve the volume of the flood equal to that of the IINGEN method but preserving the peak flow; the new hydrograph is shown in Figure 6. The factor multiplies all hydrograph ordinates except the peak flow and was of 0.8972.

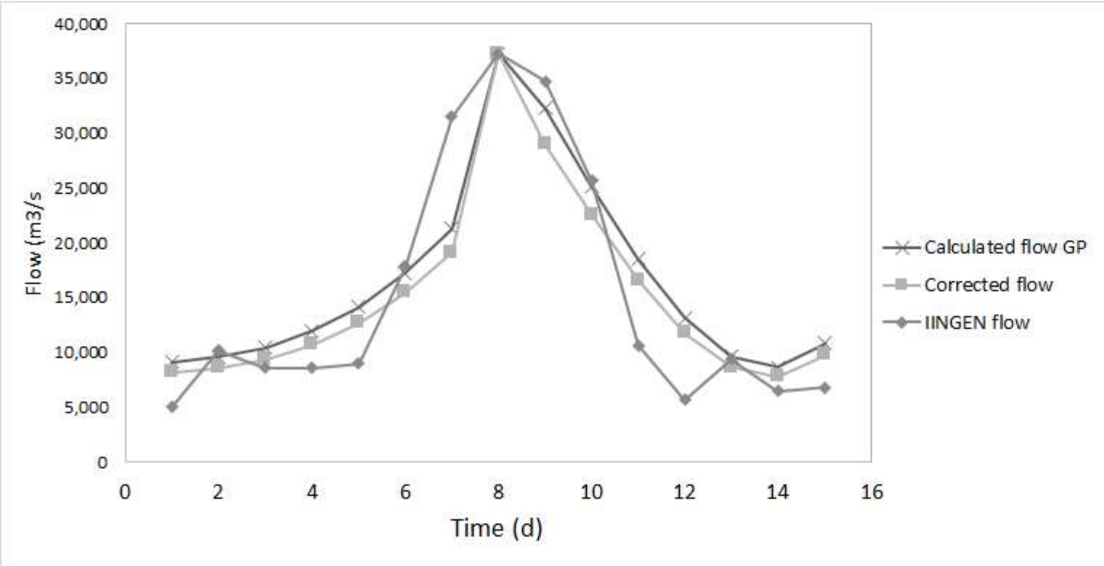

Fig. 6. GP corrected hydrograph to preserve the volume of the flood of IINGEN Method

Finally, the data of a design peak flow Q_p=*19 900* m^3/s and a design volume *V=6 835.94* millions of cubic metres obtained with a bivariate method for a return period of 10 000 years, considering different time base in the historic floods (Arganis *et al*. 2015), and the factors of GP see the Table 2 in column two were used to shape the design flood (Table 2), in the column two, then a factor of *0.5230* was obtained and applied in all ordinates except at the peak flow to adjust the volume of the flood (Figure 7). This resulted in a less robust hydrograph shape.

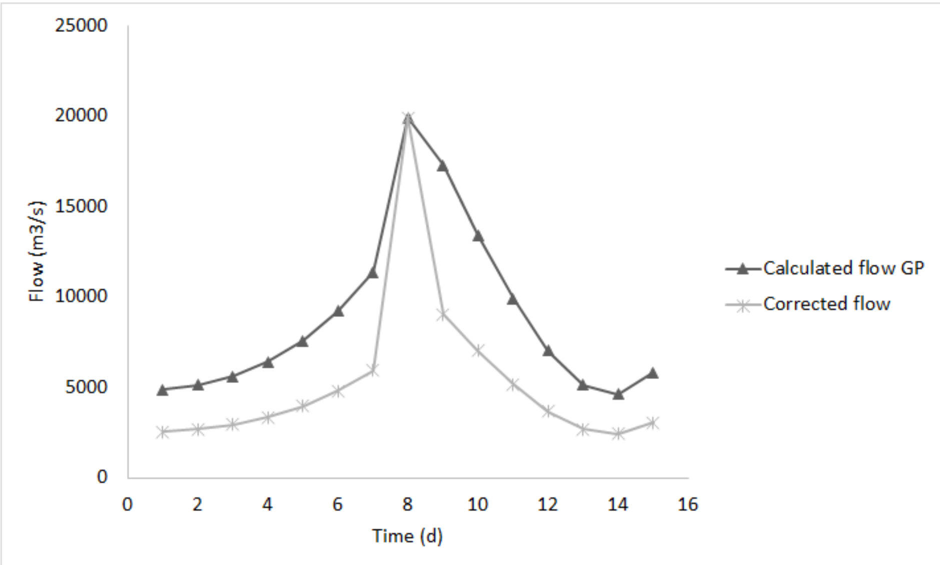

Fig. 7. Design hydrograph obtained for Tr = 10 000 years Q_p = 19 900 cubic metres per second, V = 6 835.94 millions of cubic metres (Bivariate). GP calculated and corrected volume

In the present journal, we show that with the bivariate method we can conserve the volume design and the peak flow; as well the hydrograph shapes, that present the historical behavior of any flood, are maintained. The last, is the contribution of our study. The historical shape is lost in other methods such as the Hermitian hydrographs and the Pearson hydrographs type. In both the base time is given under a triangular hydrograph, therefore the shapes always correspond to a single peak flow; meanwhile, with the genetic program method we consider the annual maximum historical floods envelope, from which is possible to get diversified shape hydrographs. Finally, with the procedure presented herein the results are nearer to the real flood's behavior "flood", so we can create design flood which simulate in a better way the profile of a analyzed dam, channels or any other hydraulic work. In an extraordinary event scenario this kind of simulation will help to transit more efficiently the flood.

4. Conclusions

A method based on applying genetic programming (GP) in the ascending and descending branch of a parameterized hydrograph was proposed in order to get a design hydrograph. Parameterization was done by considering peak flow, the peak time and the approximate time base estimated from the behavior of the maximum historical floods envelope measured in the study site.

The method of determining the shape of the ascending and descending branches of the maximum historic parameterized floods using polynomial functions obtained with GP is useful to shape the design flood similar to the historical behavior. This method has a correspondence rule which approximates the behavior of the flood and this represents the advantage respect to the increase of the annual maximum historical floods envelope. Additionally, it was possible to obtain a factor with allows corrections in the flood volume, keeping the peak flow and the time base proposed, which was useful when design peak flow and volume were obtained with bivariate methods. The method is an alternative to the IINGEN Method with the advantage that only requires analysis of maximum annual flow with one day duration and an assumed time base, reducing time calculations to shape the design flood.

References

[1] Aldama, Á. A., Ramirez, A. I. "Dam design flood estimation based on bivariate extreme-value distributions", *The extremes of the extremes: extraordinary floods: IAHS, 2002,Wallingford UK,* (271), pp. 257–262;

[2] Arganis M. L., Domínguez M. R., Rodríguez V. K. M., Preciado J., Ocón A. G., "Estimation of the bivariate distribution function of maximun historic floods", 21(2), Valencia, España, 2013, Memorias de las III Jornadas de Ingeniería del Agua, pp. 1-8;

[3] Chapra S. C., Canale R. P. "Numerical methods for engineers". Mc Graw Hill (ed.) 5 th edn, México, 2000;

[4] CONAGUA ftp://ftp.conagua.gob.mx/Bandas/Bases_Datos_Bandas (accessed 26 August 2016);

[5] Danandeh M. A., Kahya E., Olyaie E. "Streamflow prediction using linear genetic programming in comparison with a neuro-wavelet technique". 505(10), 2013,Journal of Hydrology, pp. 240-249;

[6] Domínguez M R., Arganis J. M. L. "Validation of methods to estimate flow rates for design discharge dam spillways with large regulating capacity". Hydrological Sciences Journal, 2012, 57(3), pp. 1-19;

[7] Domínguez M. R., Arganis J. M. L., Carrizosa E. E., Ramírez M. R., Gómez G. F. "Hydrologic study and determination of the operating policies of the vessels storage dams cascading over the Balsas River Balsas: El Caracol (CH Carlos Ramírez Ulloa), Adolfo Lopez Mateos (CH Infiernillo) and Jose Ma. Morelos y Pavon (CH La Villita))", Report CFE, Federal Electricity Commission, Mexico;

[8] Dorado J., Rabuñal J. R., Pazos A., Rivero D., Santos A., Puertas J. "Prediction and modeling of the rainfall-runoff transformation of a typical urban basin using ANN and GP". Applied Artificial Intelligence, 17(4), 2003, pp. 329-343;

[9] Fuentes M. O. A., Domínguez M. R., Arganis J. M. L., Herrera A. J. L., Carrizosa E. E., Sánchez C. J. A. "Estimate of design Hydrographs for the Angostura dam, Sonora, using statistical and spectral methods". Water Resources Management. Water Resour Manage, 2015, 29(11), p.p. 4021-4043;

[10] Goldberg D. E., Holland J. H. "Genetic Algorithms and Machine Learning. Kluwer Academic Publishers", 3(2), 1988, pp. 95-99;

[11] Goldberg D. E. "Genetic algorithms in search, optimization and machine learning". Addison-Wesley (ed.), 1a edn, New York, 1989;

[12] Gómez G. F. "Update the desing floods of the spillway of the dam "El Infiernillo"". Master's thesis, Universidad Nacional Autónoma de México, Mexico, 2015;

[13] Hiemstra L. A. V., Francis D. M. "The run hydrograph. In Theory and application for flood predictions", 107(6), 1979, Pretoria: Water Research Commission, pp. 759-775.

[14] Jiménez E.M. "Integral design of weirs. Ph.D thesis, Universidad Nacional Autónoma de México, México, 2000;
[15] Koza J. R. "Genetic programming, On the Programming of Computers by Means of Natural Selection". In: A Bradford Book (ed.), 1st. edn. London, England, 1992;
[16] Luthe R .A., Olivera S. F. "Métodos numéricos", Limusa Noriega (ed.), Mexico, 1994;
[17] Meshgi A., Schmitter P., Chui-Fong T. M., Babovic V. "Development of a modular model to quantify streamflow runoff from different land use Contributions in tropical urban environments using Genetic Programming". 525, 2015, Journal of Hydrology, pp. 711-723;
[18] Nourani V., Komasi M. & Alami M. "Hybrid Wavelet-Genetic Programming Approach to Optimize ANN Modeling of Rainfall-Runoff Process". Journal of Hydrologic Engineering 17(6), 2012, pp. 724-741;
[19] Pegram G. G. S., Deacon M. P. "Extreme flood Hydrographs of Chosen probability. International Symposium on Dams and Extreme Floods". International Symposium on Dams and Extreme Floods, ICOLD. Granada, Spain 16 September 1992, 17(6);
[20] Rabuñal J. R., Puertas J., Suárez J., Rivero D. "Determination of the unit hydrograph of a typical urban basin using genetic programming and artificial neural networks". Hydrological Processes, 21(4), 2007, pp. 476-485;
[21] Ramírez A. I., Aldama A. "Frequency joint analysis for the estimation of design floods", , IMTA, Serie Avances en Hidráulica, México, 2000;
[22] Savic D. A., Walters G., Davidson J. W. "A genetic programming approach to rainfall-runoff modeling. Water Resources Management", 13(3), 1999, pp. 219-231;
[23] http://www.sc.ehu.es/sbweb/energias-renovables/index.html], (accessed 26 August 2016);
[24] Whigham P. A., Crapper P. F. "Modelling rainfall-runoff using genetic programming". 33(7), 2001, Mathematical and Computer Modelling, pp. 707-721.