APPLYING A FLEXIBLE FUZZY ADAPTIVE REGRESSION TO RUNOFF ESTIMATION

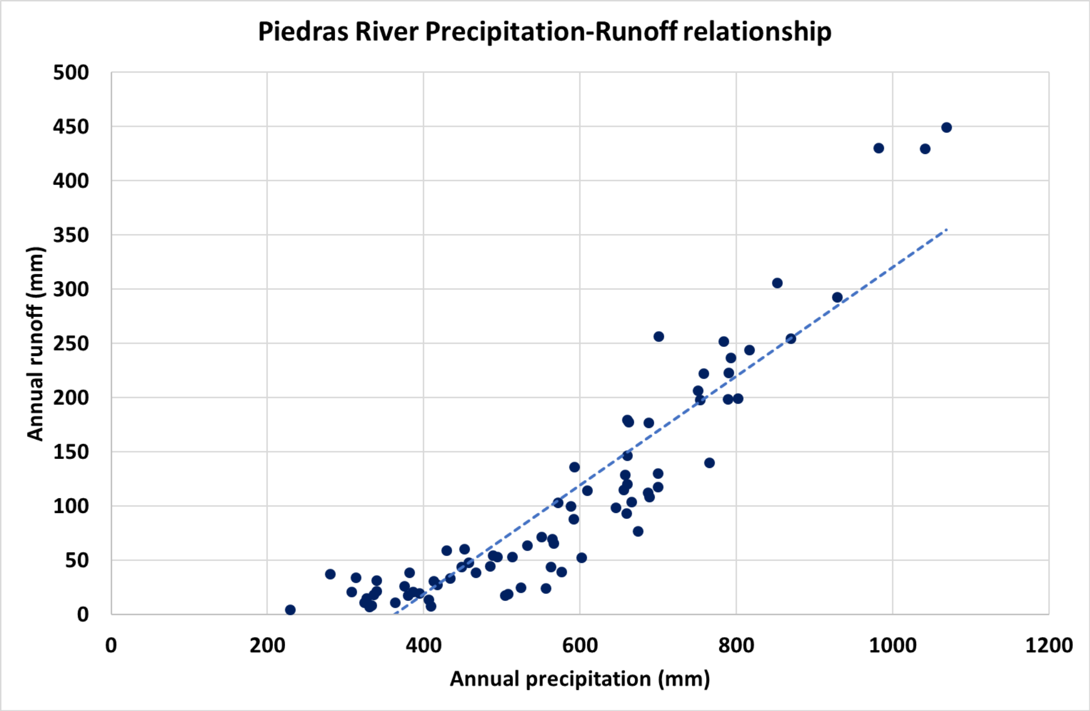
Mike SPILIOTIS[1], Luis GARROTE[2]

[1] Department of Civil Engineering, School of Engineering, Democritus University of Thrace, Greece

[2] Department of Civil Engineering: Hydraulics, Energy and Environment, Universidad Politécnica de Madrid, Spain

Scope of the work

- $E = k(P - P_0)$,LINEAR APPROACH
- However, in certain arid or semiarid climates, annual precipitation may be lower than the runoff threshold -> yield negative runoff.
- In this article, an adaptive fuzzy based regression is proposed to represent the non-constant behaviour of the **relationship** between precipitation and runoff at the annual scale:
  - For high precipitation, beyond a fuzzy threshold, a conventional (crisp) relation between precipitation and runoff is established,
  - For low precipitation, a curve with lower slope must be derived.
  - Between these curves, and for a precipitation range close to the runoff threshold, each curve holds to some degree.



FUZZY RULE BASED SYSTEM (1)

Ru(λ): If x_1 is A_1^{λ} and x_2 is A_2^{λ}……and x_N is A_N^{λ}

then $y = \alpha_{\lambda 0} + \alpha_{\lambda 1} x_1 + \ldots\ldots + \alpha_{\lambda N} x_N$

→ CRISP COEFFICIENTS MODEL 1

→ FUZZY COEFFICIENTS MODEL 2

A_n^{λ}: fuzzy set

λ: order of the rule, λ = 1, 2,…..L

L: number of the fuzzy rules

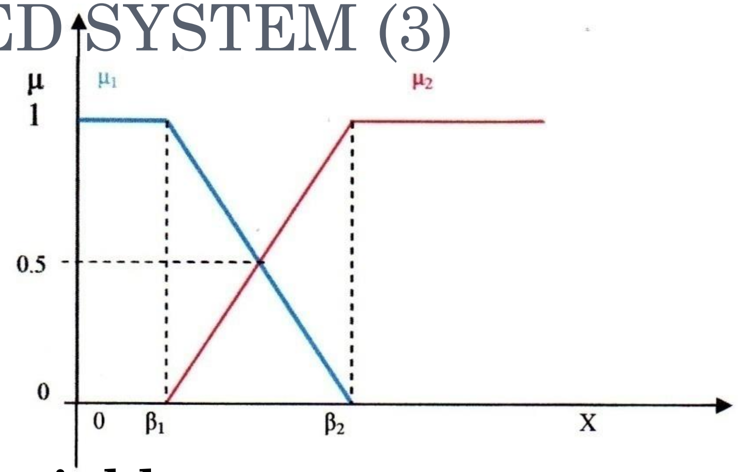
x: independent variable

N: number of the independent variables

y: dependent variable

Simplified architecture of the membership functions

FUZZY RULE BASED SYSTEM (3)



$$\mu_1(P) + \mu_2(P) = 1$$

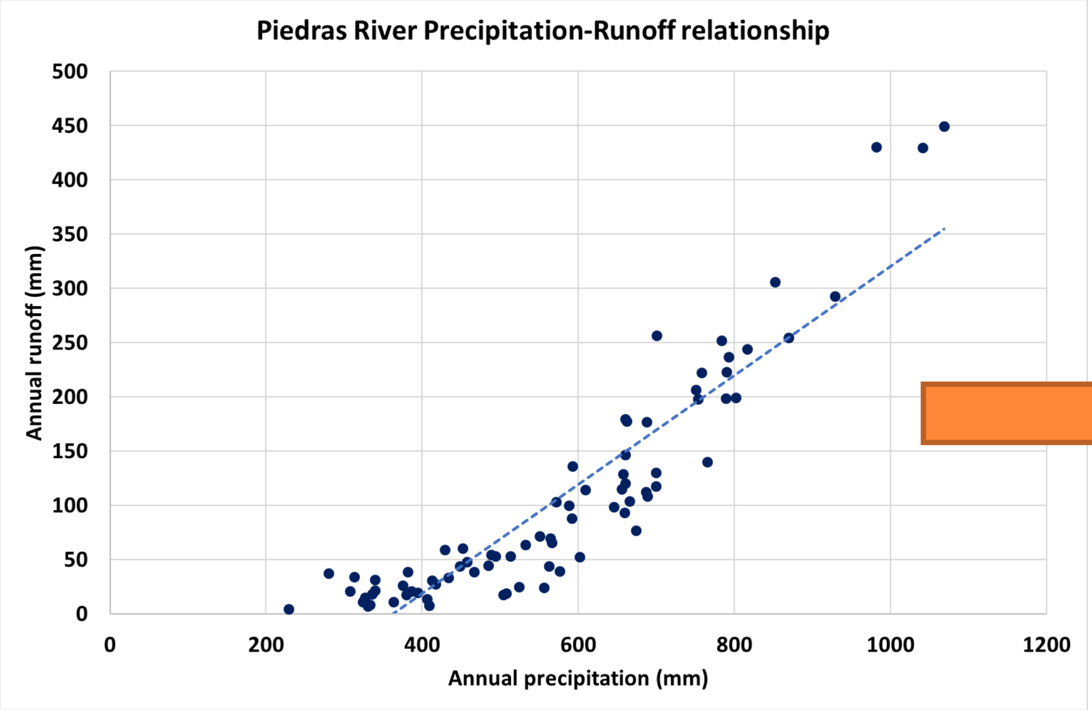
- **Two rules, one key variable**
- **Two regions without uncertainty – Two conventional regression equations (or fuzzy regressions)**
- **Between two crisp regions there is a grey (fuzzy) region where both rules are activated to some degree.**
- **Grey (fuzzy) region: between β_1 and β_2**

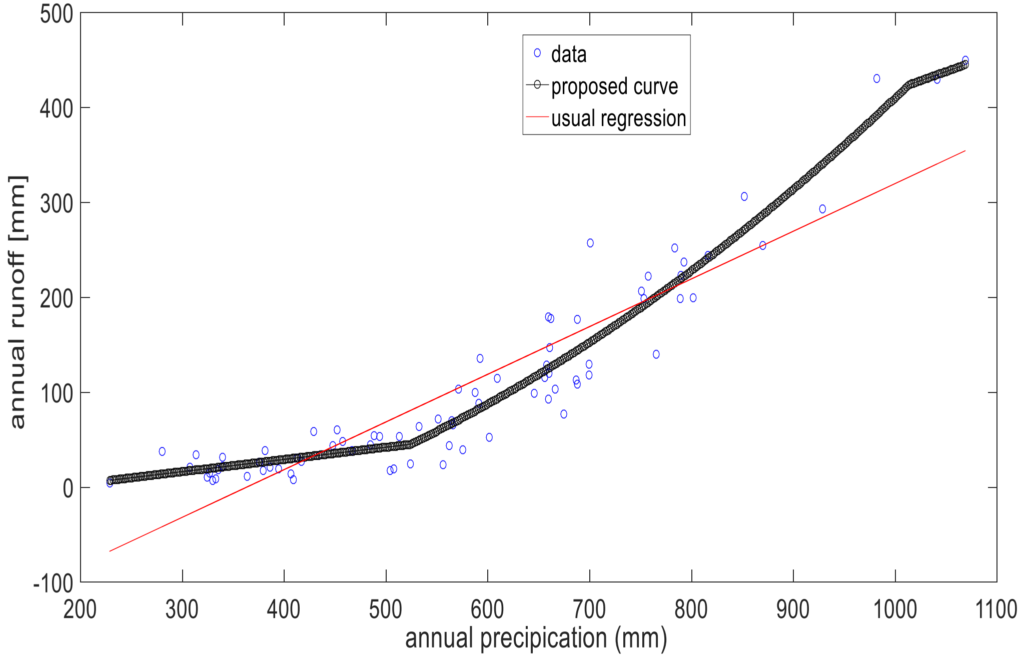
MODEL 1

IF P (annual precipitation) is low THEN y (annual Runoff) is $(y = a_{10} + a_{11}P)$

IF P (annual precipitation) is high THEN y (annual Runoff) is $(y = a_{20} + a_{21}P)$

(MODEL 1)





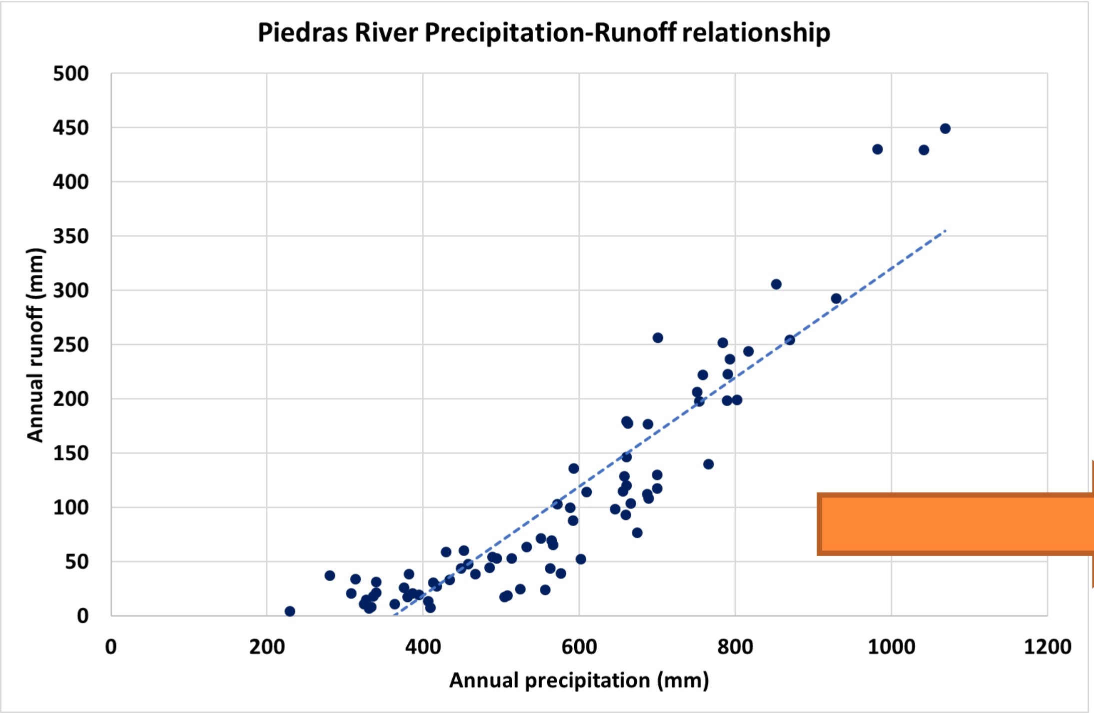
Although the fuzzy reasoning is used the output is a crisp curve

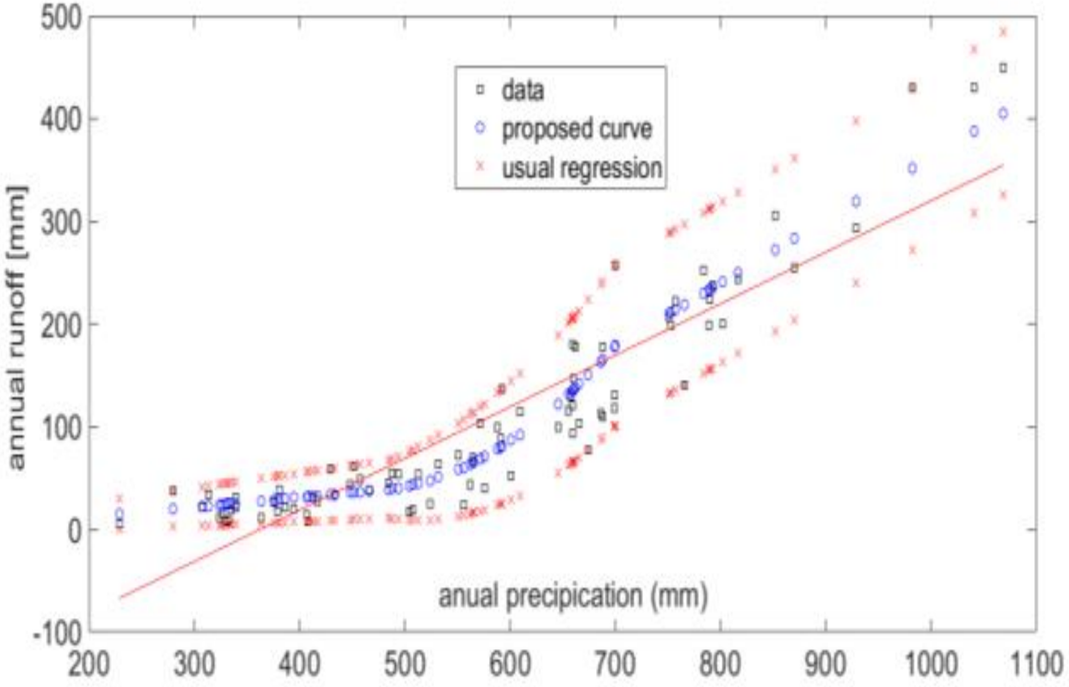
MODEL 2

IF P is low THEN Annual Runoff, yis $\left(\tilde{y} = \tilde{a}_{10} + \tilde{a}_{11}P\right)$

IF P is high THENAnnual Runoff,y is $\left(\tilde{y} = \tilde{a}_{20} + \tilde{a}_{21}P\right)$

(MODEL 2)

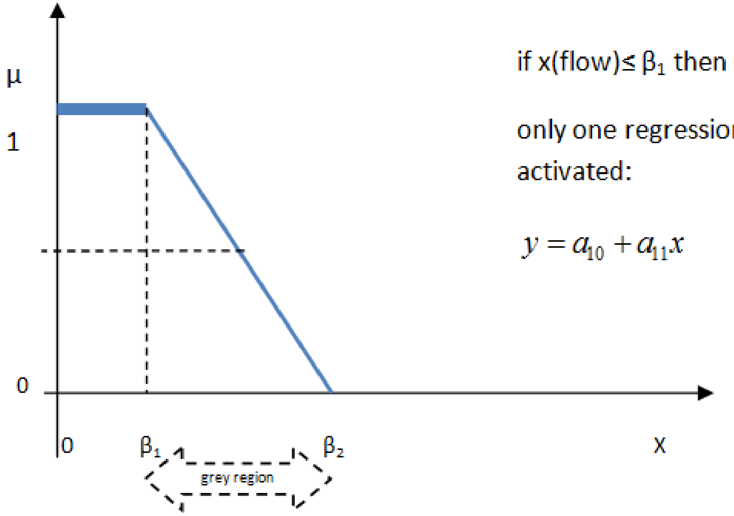




A fuzzy band is produced where all the data must be included within the produced fuzzy band

Three cases, two crisp cases model 1

First case, low precipitation

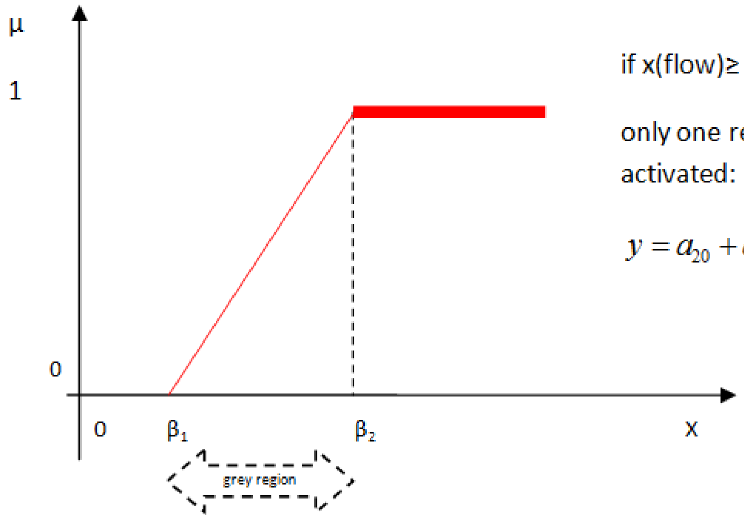


if x(flow)≤ β_1 then $\mu_1=1$, $\mu_2=0$

only one regression relation is activated:

$$y = a_{10} + a_{11}x$$

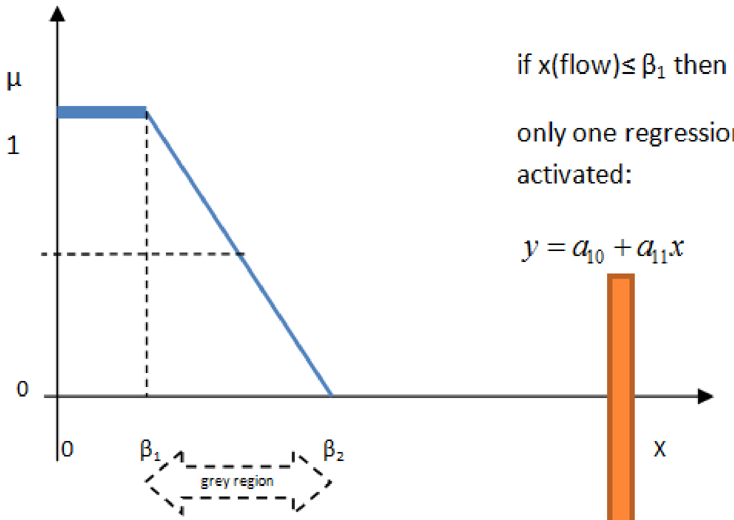
Second case, high precipitation



if x(flow)≥ β_2 then $\mu_2=1$, $\mu_1=0$

only one regression relation is activated:

$$y = a_{20} + a_{21}x$$

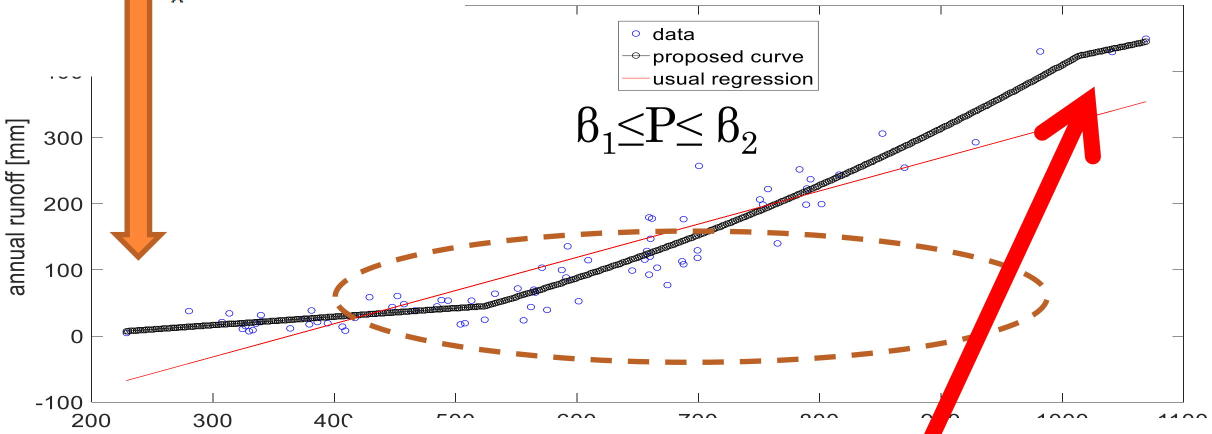


if x(flow)≤ β_1 then μ_1=1, μ_2=0

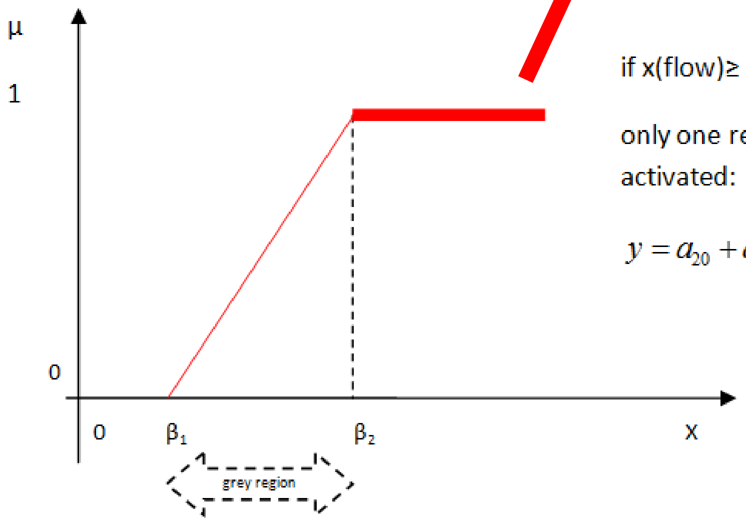
only one regression relation is activated:

$$y = a_{10} + a_{11}x$$

Grey zone



$\beta_1 \leq P \leq \beta_2$

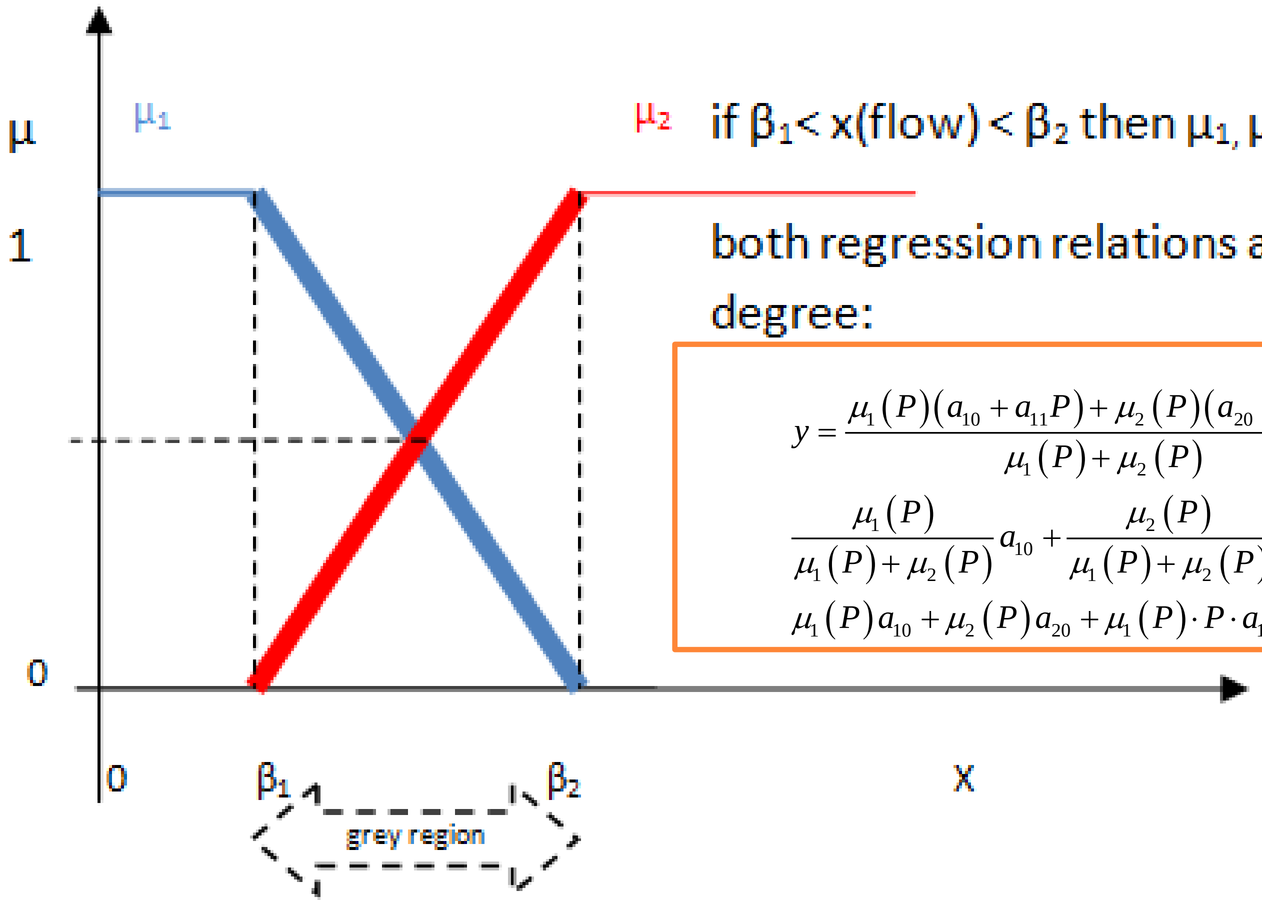


if x(flow)≥ β_2 then μ_2=1, μ_1=0

only one regression relation is activated:

$$y = a_{20} + a_{21}x$$

One grey region, 3rd case



if $\beta_1 < x(\text{flow}) < \beta_2$ then $\mu_1, \mu_2 \neq 0$

both regression relations are activated at some degree:

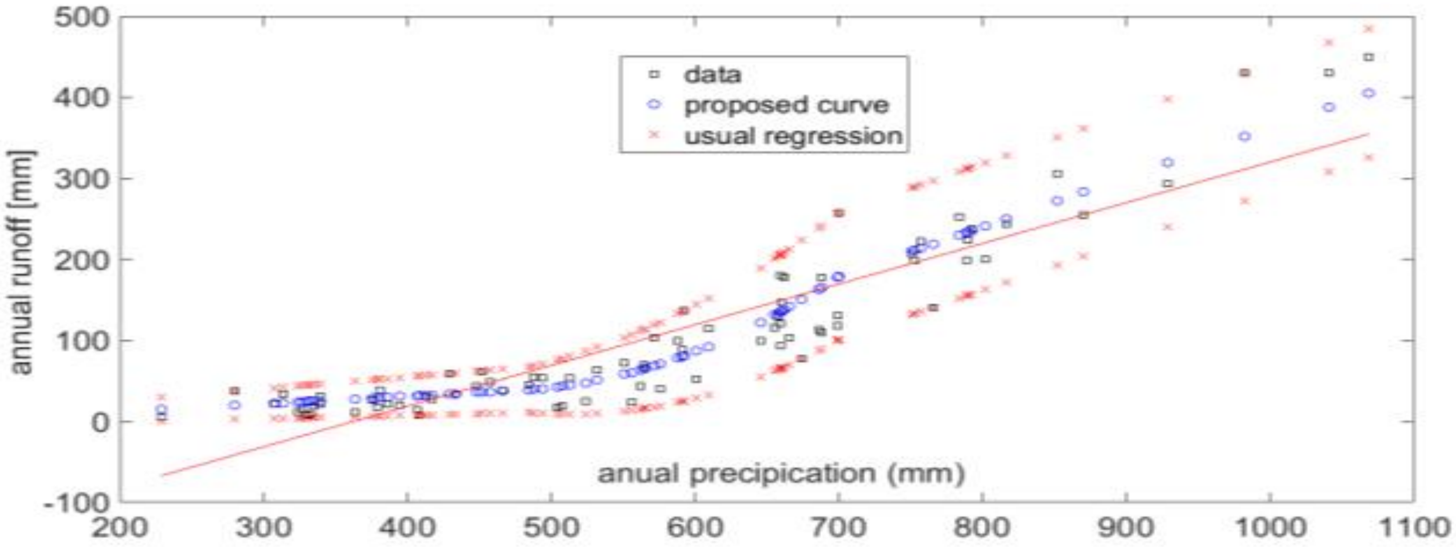
$$y = \frac{\mu_1(P)(a_{10} + a_{11}P) + \mu_2(P)(a_{20} + a_{21}P)}{\mu_1(P) + \mu_2(P)} =$$

$$\frac{\mu_1(P)}{\mu_1(P)+\mu_2(P)} a_{10} + \frac{\mu_2(P)}{\mu_1(P)+\mu_2(P)} a_{20} + \frac{\mu_1(P)\cdot P}{\mu_1(P)+\mu_2(P)} a_{11} + \frac{\mu_2(P)\cdot P}{\mu_1(P)+\mu_2(P)} a_{21} =$$

$$\mu_1(P) a_{10} + \mu_2(P) a_{20} + \mu_1(P)\cdot P \cdot a_{11} + \mu_2(P)\cdot P \cdot a_{21}$$

Kosko and Tzimopoulos: overlap at least 25%

Model 2



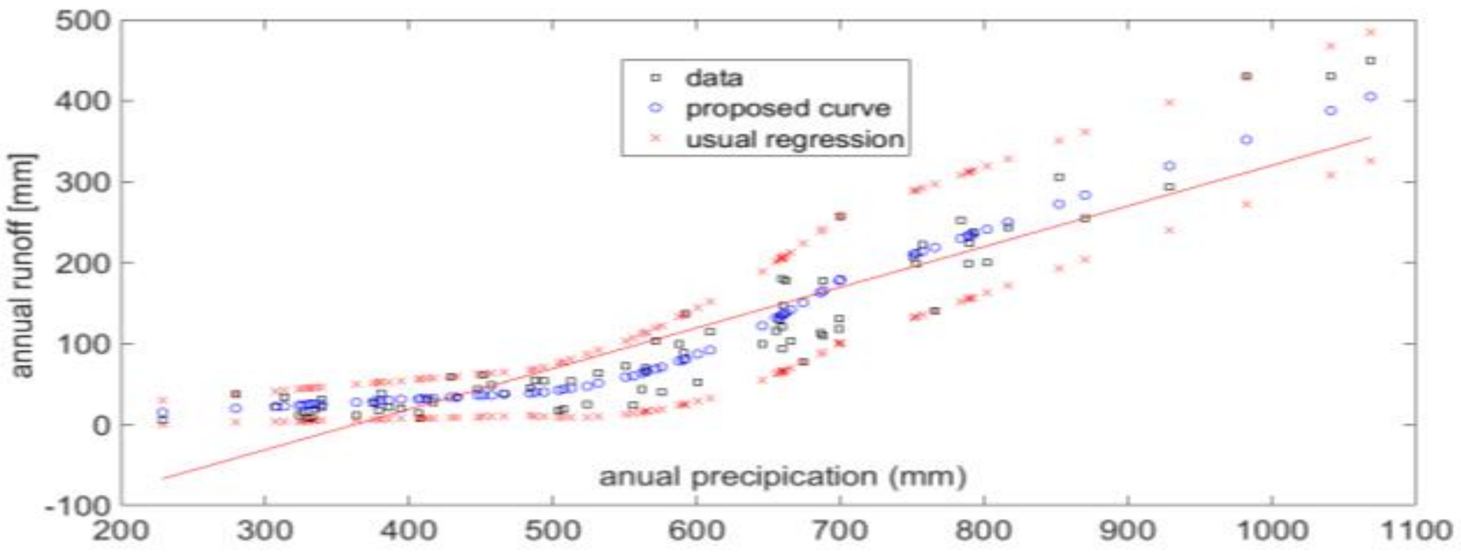
Low precipitation
Crisp region
Fuzzy regression "simple"
Fuzzy regression of Tanaka

Grey region

High precipitation
Crisp region
Fuzzy regression "simple"
Fuzzy regression of Tanaka

Model 2

All the data must be included within the produced fuzzy band
Goal:
Minimum spread of the produced fuzzy band



Grey region

In this new model, where the coefficients $(\tilde{a}_{10},\ \tilde{a}_{11},\tilde{a}_{20},\tilde{a}_{21})$ are fuzzy symmetrical triangular numbers, so a nonlinear fuzzy curve is produced. Hence, the problem concludes to the following equation:

$$y=\frac{\mu_1(P)(\tilde{a}_{10}+\tilde{a}_{11}P)+\mu_2(P)(\tilde{a}_{20}+\tilde{a}_{21}P)}{\mu_1(P)+\mu_2(P)}=$$

$$\mu_1(P)\tilde{a}_{10}+\mu_2(P)\tilde{a}_{20}+\mu_1(P)\cdot P\cdot\tilde{a}_{11}+\mu_2(P)\cdot P\cdot\tilde{a}_{21}$$

TWO MODELS

Model 1

- **Coefficients: Crisp numbers**
- **Coefficients: least squares method**
- **If the thresholds β_1 and β_2 are known, the coefficients can be determined on the basis of the least squares method**
- **PSO: A population of possible solutions (containing β_1 and β_2) is modulated.**
- **Output: Crisp number**

Model 2

- **Coefficients: Fuzzy numbers**
- **Coefficients: Minimum fuzzy band**
- **If the thresholds β_1 and β_2 are known, the coefficents can be determined. -> Concludes to a LINEAR PROGRAMMING PROBLEM**
- **PSO: A population of possible solutions (containing β_1 and β_2) is modulated.**
- **Output: Fuzzy number**

PARTICLE SWARM OPTIMIZATION METHOD (2)
PSO ALGORITHM

1. **Initialize a population array of particles with random positions and velocities on D dimensions in the search area!**
2. **Loop!**
3. **For each particle evaluate the desired optimization fitness function in D variables!**
4. **Compare particle fitness evaluation with its best previously visited position (p_i)! If the current value is better than p_i, then set p_i equal to the current value!**
5. **Identify the particle in the neighborhood with the best success so far, and assign its index to the variable p_g!**
6. **Change the velocity and position of the particle according to the following equation:**

PARTICLE SWARM OPTIMIZATION METHOD (3)
PSO ALGORITHM

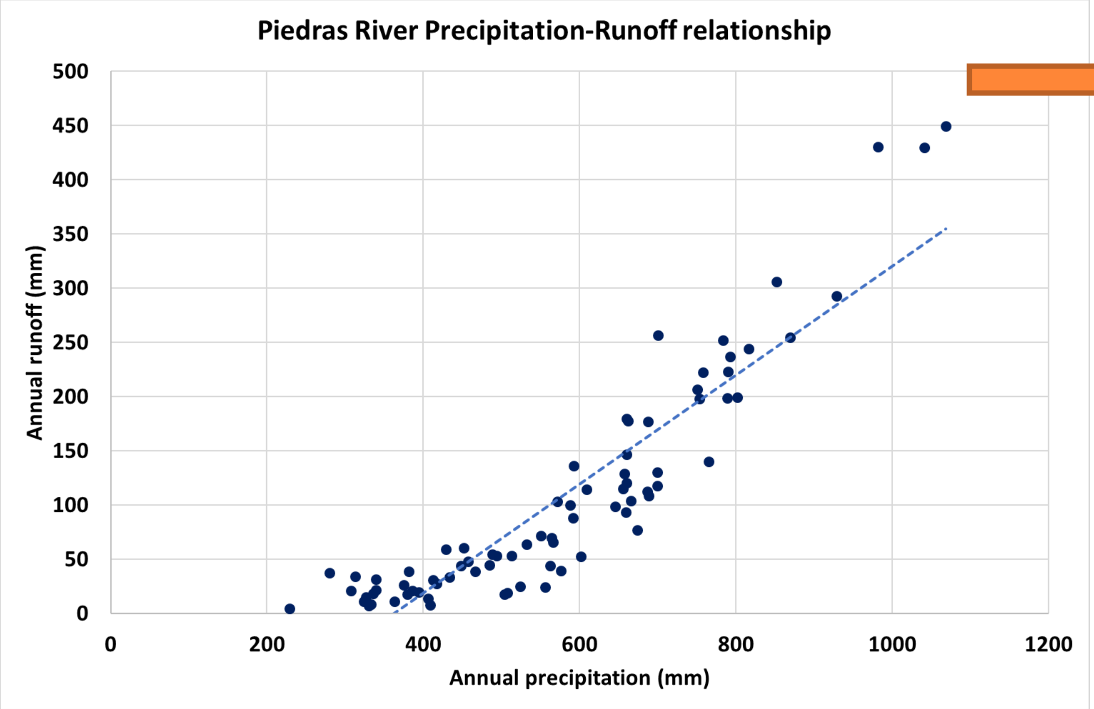
New Position

$$\begin{cases} \vec{\upsilon}_i(t+1) = \omega\vec{\upsilon}_i + c_1\rho_1(\;)\cdot\left(\vec{p}_i - \vec{x}_i(t)\right) + c_2\rho_2(\;)\cdot\left(\vec{p}_g - \vec{x}_i(t)\right) \\ \vec{x}_i(t+1) = \vec{x}_i(t) + \vec{\upsilon}_i(t+1) \end{cases}$$

individuality

sociality

CASE STUDY (1) MODEL 1



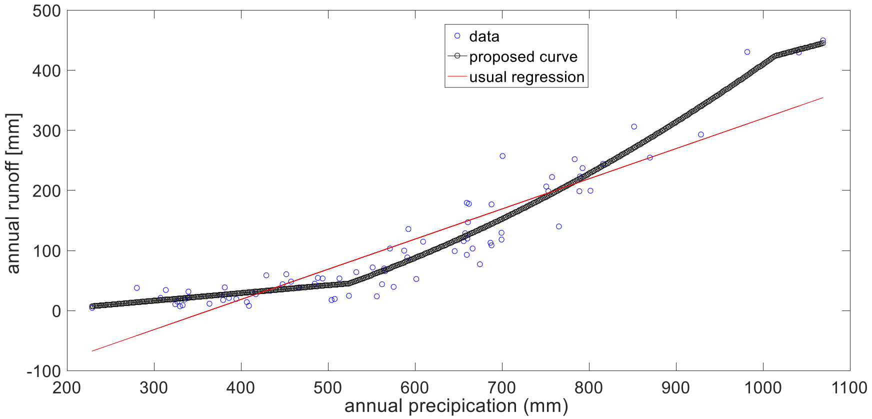


Figure 2. The proposed method and the conventional regression applied in order to assess a relation between annual precipitation and runoff in the case of Rio Piedras

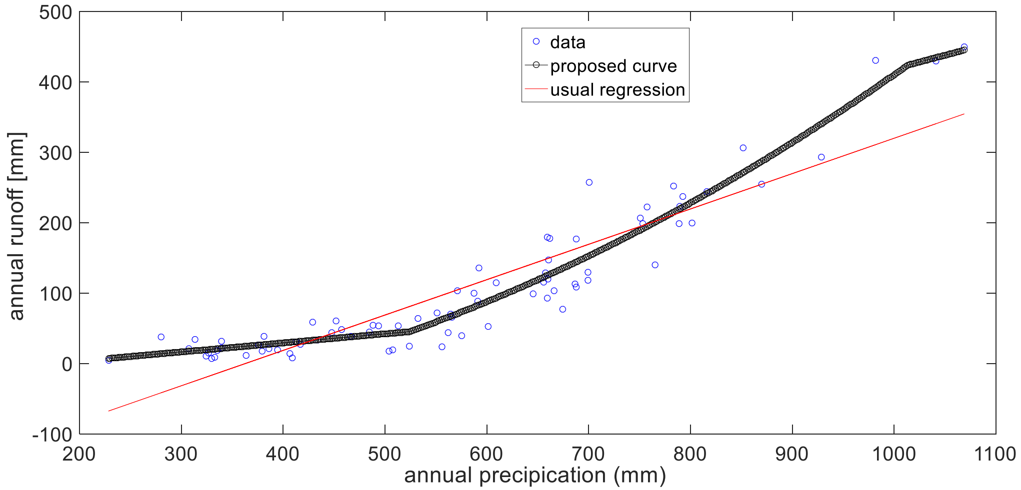
The proposed method recognizes the grey region between regarding the precipitation (524.06, 1,013.43) that is, a large non linear behavior. The produced equation is:

$$DICH(P) =$$

$$\mu_1(P)(-522.1599\ ) + \mu_2(P)(35.8763\ ) + \mu_1(P)\ 0.1287 \cdot P + \mu_2(P)0.3828 \cdot P$$

- **E′ (coefficient of efficiency Nash-Sutcliffe) = 0.95 vs 0.89 (conventional regression)**
- **Curves: expected monotonic form**

Discussion-conclusions



$$DICH(P) =$$

$$\mu_1(P)(-522.1599?) + \mu_2(P)(35.8763?) + \mu_1(P) \quad .1287 \cdot P + \mu_2(P)0.3828 \cdot P$$

Figure 2. The proposed method and the conventional regression applied in order to assess a relation between annual precipitation and runoff in the case of Rio Piedras

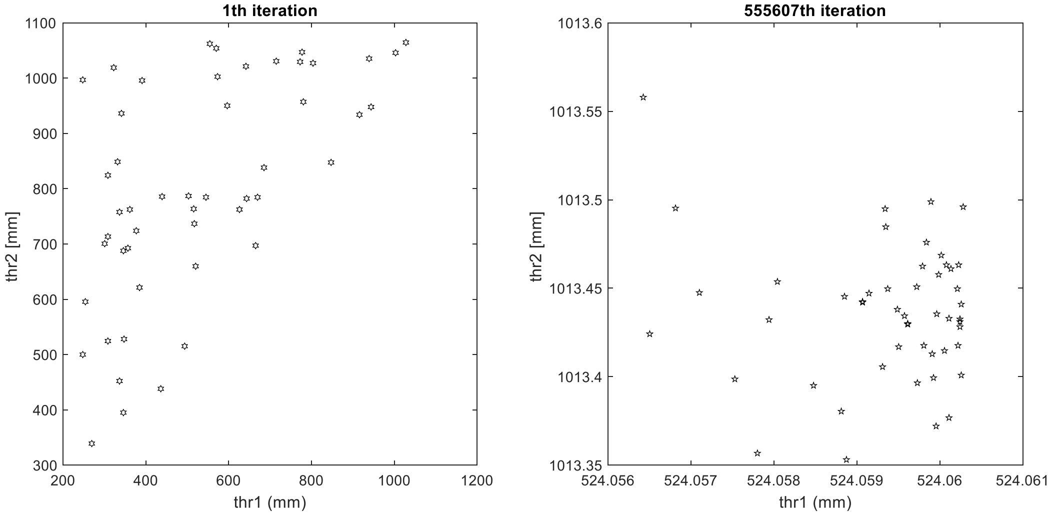
- Area**: The Piedras River** is a coastal river in the southwest of Spain. It drains a contributing basin of 550 km^2, running from north to south along 40 km in the Huelva province. Mean annual precipitation is 574 mm/yr and mean annual runoff is 106 mm/yr. It is regulated by the Piedras and Los Machos reservoirs, which are operated for water supply and irrigation.

IMPROVED RESULTS

from the weighted average of best positions of Eq. (11). The user-defined value v_j^{max} is usually equal to a fraction of the search space along the j-th coordinate direction, i.e.,

$$v_j^{max} = \lambda_j \left(x_j^{max} - x_j^{min} \right), \qquad j \in D, \ \lambda_j \in (0, 1). \qquad (13)$$

Psaropoulos in R.Martí et al. (eds.), *Handbook of Heuristics*,
DOI 10.1007/978-3-319-07153-4_22-1



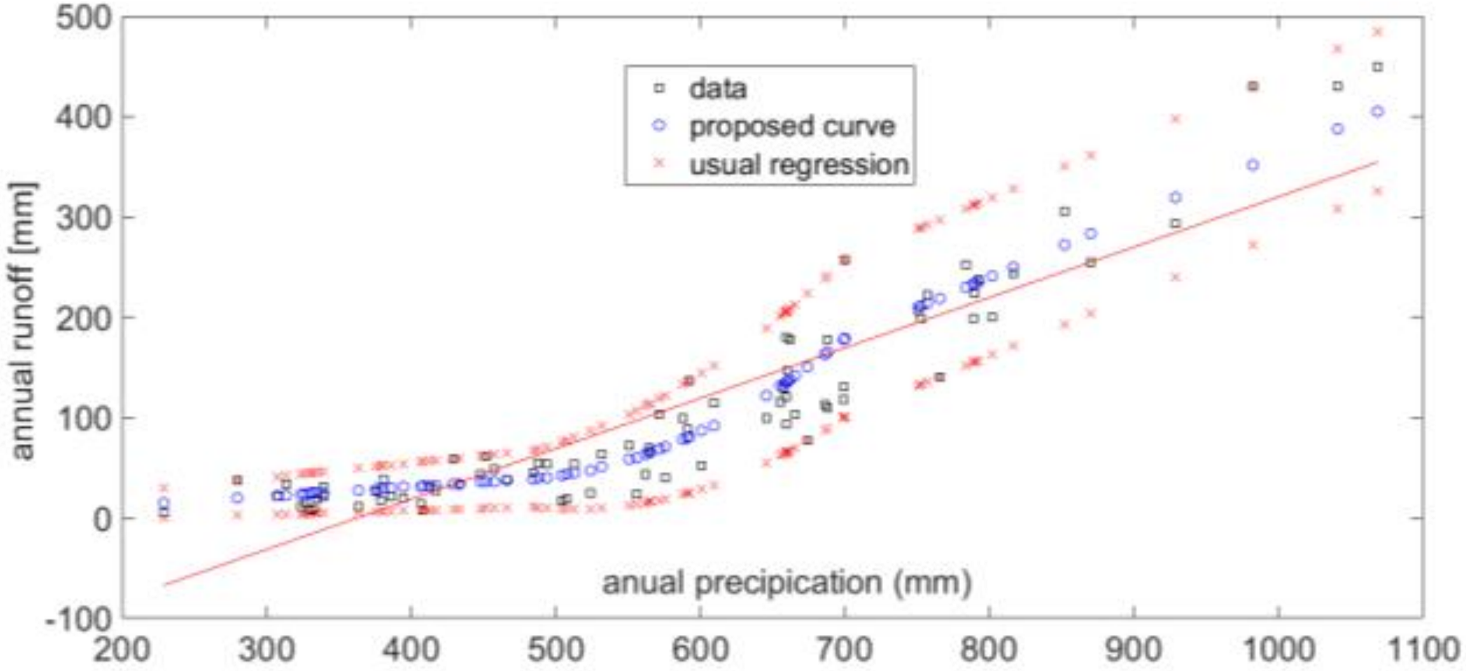
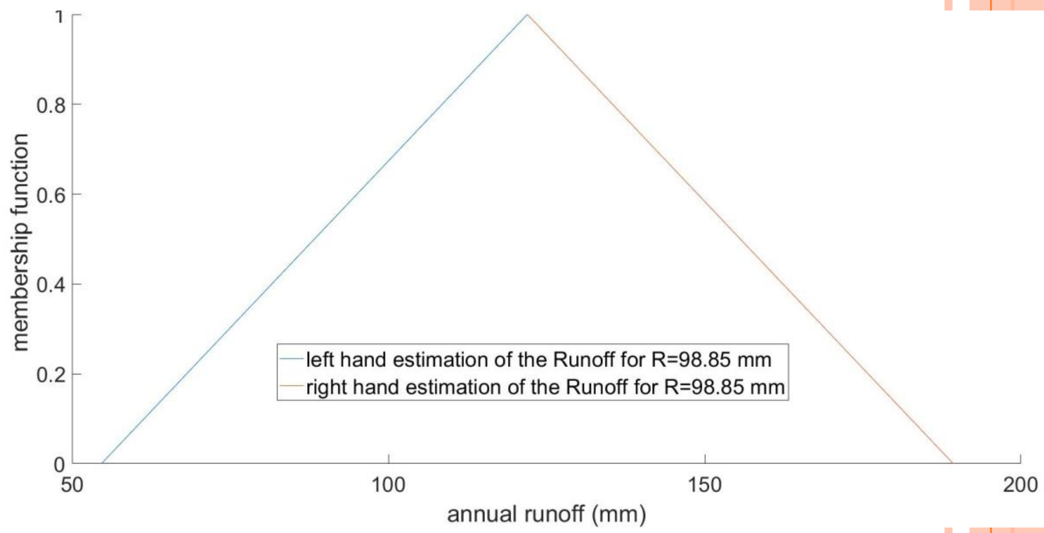


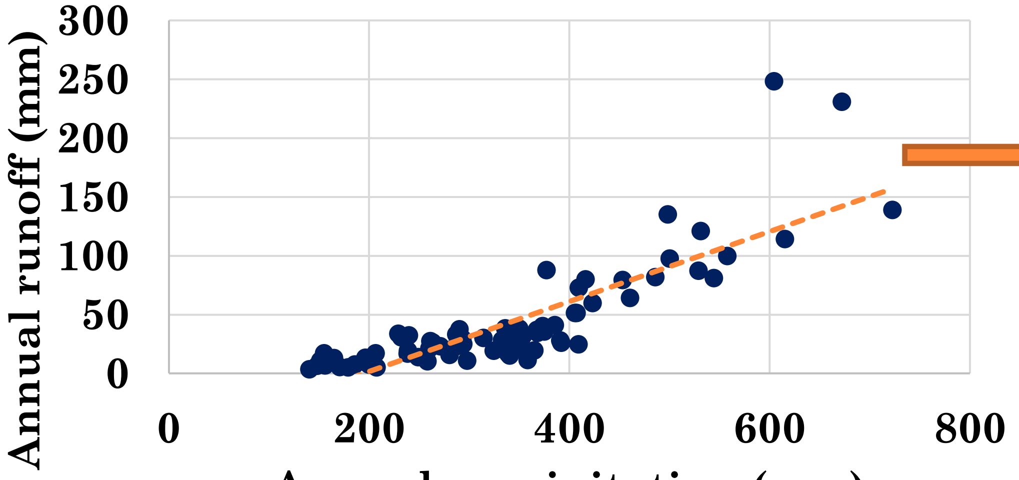
Figure 5. The proposed method with fuzzy regression curves applied in order to assess a fuzzy relation between annual precipitation and annual runoff in the case of Rio Piedras.

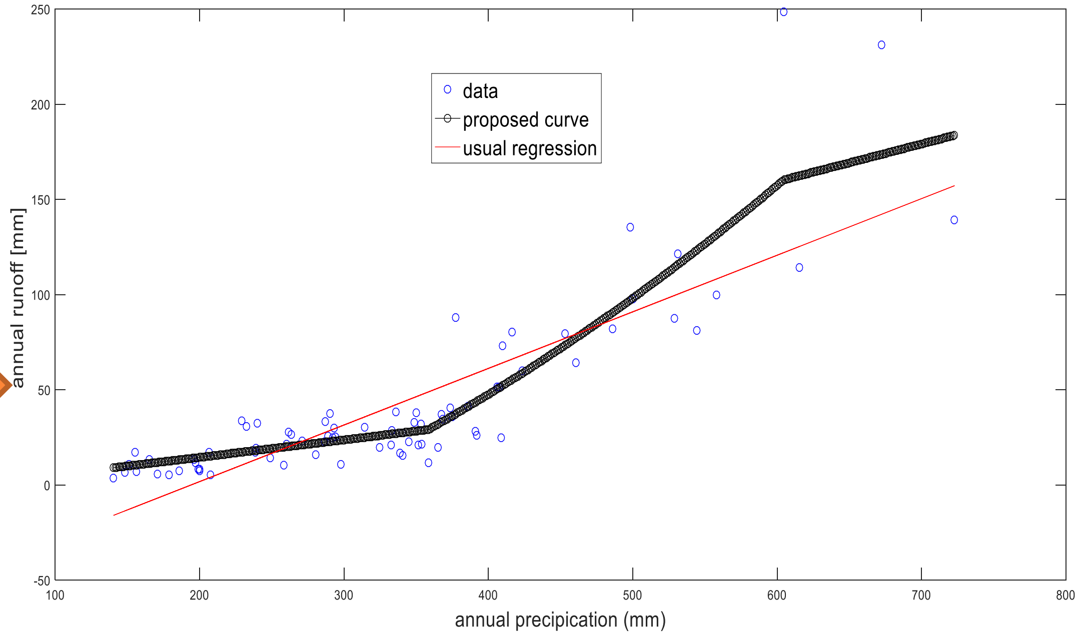
$$DICH(P) =$$
$$\mu_1(P)(-6.3182,\ 3.3334) + \mu_2(P)(\ 250.4427,\ 76.0990) + \mu_1(P)(\ .0933,\ 0.0510)P + \mu_2(P)(\ .6130, 0.0029)\cdot P$$

All the data must be included within the produced fuzzy band



Aguas River Precipitation-Runoff relationship





The Aguas river is a short coastal river in the south of Spain, running along 65 km through the east of the province of Almería. The contributing basin is 547 km^2. The climate is semiarid, with mean annual precipitation of 334 mm/yr and mean annual runoff of 41 mm/yr.

Conclusion

- Two hybrid methods with no linear behavior
- The main difference between model 1 and models 2 is that in the last model all the data must be included within the produced fuzzy estimation of the annual runoff.
- The proposed method is suitable to estimate the annual water yield, but it is not intended to assess the peak flow under a significant rainfall. However, the second model requires more computational time.
- with our approach, we are able to describe the Precip-Runoff behavior precisely in the most critical years, with minimum precipitation. This is very relevant for drought management.

Xanthi

Thanks for your attention!




Madrid