MedCrave
Step into the World of Research

International Journal of Hydrology

Research Article

Open Access

CrossMark

Calculation method for the classification of water basins shape within a specific basins set

Abstract

This paper relies on concepts from analytical geometry, differential and integral calculus, and statistics to develop a calculation script to analyze watersheds vis-à-vis their shapes. Said script consists of determining whether or not each watershed studied has a shape the deviates from the average behavior of the set. For this purpose, the method considers the approximation of the shape of each basin by ellipses, so that the semi-axes of the determined ellipses are the input parameters of the tool. The sequence of equations to be adopted is based on the existence of a trend line that reflects the average behavior of the set, while each scattered point corresponds to the shape condition of a basin. Applying the method to a group of neighboring basins located in Serra da Mantiqueira showed that the projections can produce different results, and also demonstrated that the script could be used with relative practicality.

Keywords: average behavior, ellipses, trend line, projections, method application

Volume 9 Issue 4 - 2025

Guilherme Cavallari Bueno, Carlos Eugenio Pereira

Faculdade de Engenharia Civil, Universidade Federal de Uberlândia, Brazil

Correspondence: Carlos Eugenio Pereira, Faculdade de Engenharia Civil, Universidade Federal de Uberlândia, Brazil, Fax +553432394170, Tel +5534991505624

Received: November 10, 2025 | **Published:** December 24, 2025

Introduction

Discussions on climate change are growing and many studies indicate that in the southeast region of Brazil precipitation will be more intense. Thus, flooding-related issues tend to increase, especially in urban areas. Studying watersheds is essential for understanding the behavior of precipitation and natural watercourses. This has an influence on the execution of Civil Engineering works concerning the construction of dams, regardless of their purposes, and on the planning of urban areas, since the rainwater harvesting system and the traffic routes of a city require prior knowledge of the water flow direction and critical flooding areas. In addition, the basin analysis is an essential factor for managing water resources, since Brazil is a country in which, in addition to essential daily activities, energy production is related to these resources.[1] In Santos et al.,[2] the morphometric indices of the Jaguaribe River basin were evaluated to confirm flooding likelihood. This basin, like many others, is undergoing an urbanization process, which causes serious environmental problems due to the silting up of river channels due to erosion processes and increased surface runoff. Along the same lines, Soares and Galvíncio[3] used the LiDAR (Light Detection and Ranging) sensor with a spatial resolution of 5 m and a scale of 1:5000 in a study carried out in the Beberibe River watershed. Based on morphometric parameters, they concluded that this basin is not naturally susceptible to flooding, but that, despite this, other factors lead to this type of problem.

Alves et al.[4] evaluated the use of large-scale models to assess flood risk areas at local, regional, and national levels in three municipalities in the State of Rio Grande do Sul. They considered precipitations with specific periods and compared the results achieved with the flood spots obtained by the Geological Survey of Brazil (CPRM), with satisfactory results. The importance of this type of study is noted in the definition of possible flooded areas, especially near urban regions. Thus, it is known that whatersheds with a shape similar to a circle are more likely to have flood regions, due to the greater difficulty of water flow towards the outlet. Therefore, the calculation method developed in this work aims to explore this proximity in a simple and efficient way, which can be measured through numerical values, such as the circularity index. Lima and Lima[5] consider this index the ratio between the watershed and circle areas. To find the area of a circle, we must know the basin perimeter, which must be equal to the circle's circumference. This allows us to know its radius and, consequently, its area. Concerning the indexes, the derivative is one of the mathematical tools that permeate the development of data analysis methods. The data analysis is essential for predicting the moment and the ideal way to carry out specific actions so that the measures taken allow the best solution for a practical problem, or even the best possible understanding of reality.[6]

Therefore, since this study is closely related to data analysis, it is also aligned with statistics. Costa Neto[7] defined it as a science that deals with the means of achieving goals, and not with the purpose itself, serving to offer information that will serve as a guide for decision making based on data and facts. Lastly, this study region is located in Serra da Mantiqueira, a mountain range located on the borders between three Brazilian states in the southeast region: Minas Gerais, Rio de Janeiro, and São Paulo.[8] This paper aims to establish a method capable of analyzing and comparing the various basins of a region regarding the degree of proximity to a circle that presents each of them. Therefore, the study aims at defining a calculation method based on simple input parameters capable of providing analytical results that can be easy to interpret. To this end, the script was applied through calculations made for a set of three basins in a given region of Serra da Mantiqueira.

Material and methods

Study area

The three hydrographic basins of this study are in Serra da Mantiqueira, a mountain range in Southeastern Brazil, close to the border with Rio de Janeiro. As shown in Figure 1, developed with the QGIS software, said basins comprise the Congonhal (blue), Piedade (brown), and Toca (pink) streams. The approximate coordinates of the outlets are -22° 08' 19.03" (latitude) and -44° 24' 39.17 (longitude) for the Congonhal stream basin, -22° 04' 30.68" (latitude) and -44° 21' 8.46" (longitude) for the Piedade stream basin, and -22° 13' 14.02" (latitude) and -44 ° 22' 59.01" (longitude) for the Toca stream basin. Santos[8] lists some of the main physical aspects of Serra da Mantiqueira: the oldest part of its geological and geomorphological structure originates from the cooling of the Earth in the Precambrian period, about more than 2 billion years ago. Over time, the rocks formed at that time suffered from the chemical action caused by

their contact with water, becoming earth hills and producing the geomorphological basis of what Ab'Sáber[9] refers to as 'seas of hills'. In addition, the entire mountain range is located in the intertropical thermal zone and at high altitudes for the Brazilian relief, with a high altitude tropical climate, hot and rainy summers, and cold, dry winters.[10] From 1961 to 1990, rainfall above 350 mm was recorded in January, one of the wettest months, and below 40 mm in September, one of the driest. In the same period, temperatures below 16º C were recorded in June, one of the coldest months, and above 22º C in December, one of the hottest.[11] Finally, the area comprises tropical forests of Atlantic Forest and associated ecosystems, with prevalence of semideciduous seasonal forests (SISEMA (Sistema Estadual de Meio Ambiente e Recursos Hidricos, 2019)[12] and IBGE (Instituto Brasileiro de Geografia e Estatística, 2019).[13]

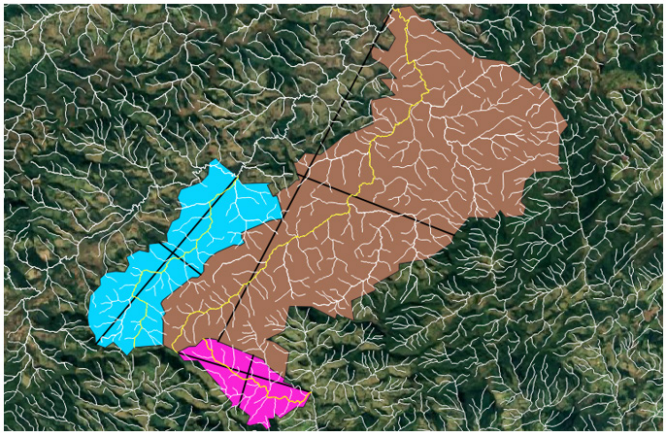

Figure 1 Watershed delineation. Source: The authors (2023).

Deduction of the fundamental equations of the method

According to Figure 2, which is consistent with the geometric and elliptical elements discussed by Winterle[14] the equation 1 can be obtained from the Pythagorean theorem:[15]

$$b^2 + c^2 = a^2 \quad (1)$$

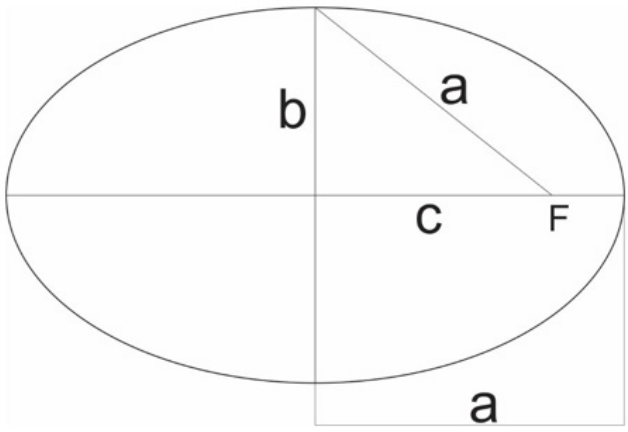

Figure 2 Ellipsis. Source: The authors (2023).

Where c is the distance between the center of the ellipsis and one of the foci, a is the greatest distance between the center and the contour line, and b is the smallest distance between the center and the contour line.

Thus, Equation 2 is valid:

$$c = \sqrt{a^2 - b^2} \quad (2)$$

Then, equations 3 and 4 below are validated:

$$\gamma = a + c = a + \sqrt{a^2 - b^2} \quad (3)$$

Where y is the greatest distance between the foci and the contour line.

$$\rho = a - c = a - \sqrt{a^2 - b^2} \quad (4)$$

Where ρ is the shortest distance between each focus and the contour line.

In order to obtain the first and second main equations, it is necessary to carry out the development of equation 5 below:

$$(b+c)^2 = b^2 + 2*b*c + c^2 \quad (5)$$

Since S the area of the right triangle enclosed in the ellipsis, equation 6 is true:

$$\frac{b*c}{2} = S \quad (6)$$

Therefore, equation 7 is also valid:

$$2*b*c = 4*S \quad (7)$$

Since p is the semiperimeter of the right triangle, we can insert equation 8:

$$p = \frac{a+b+c}{2} \quad (8)$$

Applying equations 1 and 7 in 5, we obtain the equation 9:

$$4*S = (b+c)^2 - a^2 \quad (9)$$

Equation 9 can be analyzed as formula 10:

$$4*S = (b+c+a)*(b+c-a) = 2*p*(2*p-2*a) = 4*p*(p-a) \quad (10)$$

So, the equation 11 is true:

$$S = p*(p-a) \quad (11)$$

Therefore, we obtain formula 12:

$$p^2 - p*a - S = 0 \quad (12)$$

Thus, assuming p as a variable, one of the solutions to this equation is p, the semiperimeter. For organization purposes, this solution will be called k, such that $k = p$.

Since j is the other solution, through the concept of sum and product for 2nd-degree equations,[16] the equation 13 is true:

$$k + j = -(-a) = a \quad (13)$$

After that, we obtain equations 14 and 15:

$$k = \frac{a+b+c}{2} \quad (14)$$

$$j = \frac{a-b-c}{2} \quad (15)$$

Based on equations 2 and 3, equations 16 and 17 are true:

$$a = \frac{\gamma+\rho}{2} \quad (16)$$

$$c = \frac{\gamma-\rho}{2} \quad (17)$$

Citation: Bueno GC, Pereira CE. Calculation method for the classification of water basins shape within a specific basins set. *Int J Hydro.* 2025;9(4):180–187. DOI: 10.15406/ijh.2025.09.00419

If equation 18 is true:

$$b^2 = a^2 - c^2 = (a+c)*(a-c) \quad (18)$$

What is proposed in equation 19 is also true:

$$b = \sqrt{(a+c)*(a-c)} \quad (19)$$

Applying equations 16 and 17 in equation 19, we obtain equation 20:

$$b = \sqrt{\gamma * \rho} \quad (20)$$

Substituting 16, 17, and 20 into 14 and 15 gives equations 21 and 22:

$$k = \frac{1}{2} * \left(\gamma + \sqrt{\gamma * \rho}\right) \quad (21)$$

$$j = \frac{1}{2} * \left(\rho - \sqrt{\gamma * \rho}\right) \quad (22)$$

Since r_c is the radius of any circle, it is possible to verify that the semi-axes *a* and *b* are equal to each other and also equal to r_c, such that c is equal to zero, since there is no distance between the center and foci in a circle, and both semi-axes are equal to the radius. Therefore, we obtain equations 23 and 24, through equations 14 and 15, for circles:

$$k = r_c \quad (23)$$

$$j = 0 \quad (24)$$

Thus, we can conclude that equations 21 and 22 are, respectively, the first and second main equations of the method, since *k* is the radius of a hypothetical circle to which the analyzed ellipsis approaches, while $j = 0$ means that the ellipsis is equal to that circle. It is important to emphasize that $j \leq 0$. Now, consider a cylinder with a circular base, whose base radius equals *r*. The plane π divides the cylinder transversely. Said plane rotates θ degrees (such that $0^o \leq \theta \leq 90^o$) around the axis of the diameter of the circle of radius *r*, described when intersecting at $\theta = 0^o$, as shown in Figure 3:

This method requires the delimitation of the hydrographic basins in a flat figure, such as a planialtimetric chart. The plane π, depicted in the figure above, is not the same plane on which each watershed analyzed is delimited, but otherwise the plane whose rotation describes the necessary angle of inclination of the intersection between the cylinder with a circular base and the plane itself (π) so that this (the intersection) results in the ellipsis whose ratio between the semi-axes $\frac{a}{b}$, discussed below, is equal to the same ratio $\frac{a}{b}$ calculated for the ellipsis that approximates the shape of the basin.

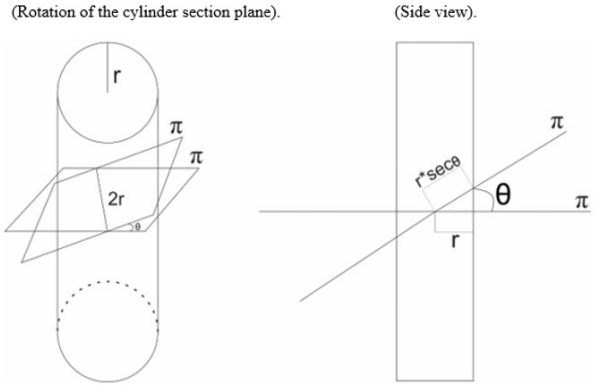

Figure 3 Cylinder and transverse intersection by plane π. Source: The Authors (2023).

As the plane π is rotated, the intersections with the cylinder begin to describe ellipses until, when $\theta = 90^o$, the longitudinal section will be an infinite rectangle. Therefore, the sections obtained as a function of the angle θ must be carried out, according to Figure 4:

Hence it is concluded that equation 25 is true:

$$\frac{a}{b} = \frac{r * sec\theta}{r} = sec\theta \quad (25)$$

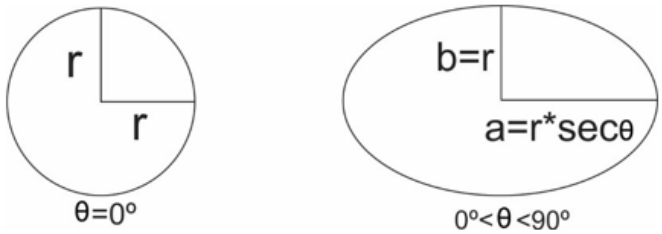

Figure 4 Sections on cylinder as a function of angle θ. Source: The authors (2023).

Where *r* is the radius of the circle obtained from the intersection of the π plane with the cylinder when $\theta = 0^o$, and θ the angle of rotation of the π plane around the axis of the diameter of the circle generated when $\theta = 0^o$.

We can determine the ab ratio for the ellipses that approximate the shape of the watersheds by transforming the referred geometric shapes into points on the curve of the $sec\theta$ function within the $0^o \leq \theta \leq 90^o$ interval. Thus, we can draw the tangent lines to the curve of the function at each point corresponding to the ellipsis of a watershed and calculate the angle formed between them. This is a parameter that indicates the behavior of the curve between the points and, therefore, a comparison between them within the function. Figure 5 shows this condition.

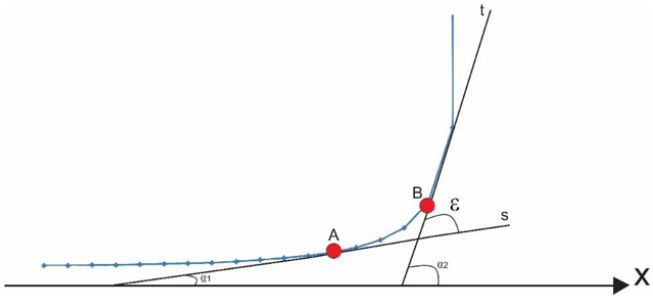

Figure 5 Tangent lines to the secant curve. Source: The authors (2023).

Since line *s* is tangent to the curve at point *A*, and line *t* is tangent to point *B*, we want to measure the angle ε, formed between them. A priori, formula 26 is true:

$$\varepsilon = \alpha_2 - \alpha_1 \quad (26)$$

Being α_1 and α_2 the angles of inclination of the tangent lines, such that always $\alpha_2 > \alpha_1$. For $sec'\theta$ as the first derivative of the function $f(\theta) = sec'\theta$, equation 27 is true:

$$sec'\theta = sec\theta * tg\theta = \frac{sen\theta}{(cos\theta)^2} \quad (27)$$

Equation 28 is, therefore, verifiable:

$$(cos\theta)^2 = \left(\frac{b}{a}\right)^2 \quad (28)$$

As the fundamental equation of trigonometry, or equation 29, states:

Citation: Bueno GC, Pereira CE. Calculation method for the classification of water basins shape within a specific basins set. *Int J Hydro.* 2025;9(4):180–187. DOI: 10.15406/ijh.2025.09.00419

$$(sen\theta)^2 + (cos\theta)^2 = 1 \quad (29)$$

Equation 30 is true:

$$sen\theta = \sqrt{1-(cos\theta)^2} = \sqrt{1-\left(\frac{b}{a}\right)^2} \quad (30)$$

Therefore, equation 31 is equally verifiable:

$$sec'\theta = \frac{\sqrt{1-\left(\frac{b}{a}\right)^2}}{\left(\frac{b}{a}\right)^2} = \left(\frac{a}{b}\right)^2 * \sqrt{1-\left(\frac{b}{a}\right)^2} \quad (31)$$

Thus, equation 32 is true:

$$\mu = tg\alpha = sec'\theta = \left(\frac{\gamma+\rho}{2*\sqrt{\gamma*\rho}}\right)^2 * \sqrt{1-\left(\frac{2*\sqrt{\gamma*\rho}}{\gamma+\rho}\right)^2} \quad (32)$$

The value of the tangent of the angle α (angular coefficient of the tangent line to the curve at the point, that is, the derivative of the function at that point) can be $\alpha = \alpha_1$ or $\alpha = \alpha_2$, and μ is the value of the tangent of the angle α, and γ' and ρ' are the parameters for the main watershed (the one in which the river with the largest volume is located). Thus, equation 32 is broken down into equation 33:

$$\mu' = tg\alpha' = sec'\theta = \left(\frac{\gamma'+\rho'}{2*\sqrt{\gamma'*\rho'}}\right)^2 * \sqrt{1-\left(\frac{2*\sqrt{\gamma'*\rho'}}{\gamma'+\rho'}\right)^2} \quad (33)$$

Can be $\alpha' = \alpha_1$ or $\alpha' = \alpha_2$, and μ' are the tangent value of the angle α'. Thus, whereas for function $sec\theta$ ($0^o \le \theta \le 90^o$), we can conclude that the greater $tg\alpha$ is, the greater α will be, and therefore $tg\alpha_2 > tg\alpha_1$. Thus, if $tg\alpha > tg\alpha'$, $\alpha = \alpha_2$ and $\alpha' = \alpha_1$, and if $tg\alpha' > tg\alpha$, $\alpha = \alpha_1$ and $\alpha' = \alpha_2$, emphasizing that $\alpha_2 > \alpha_1$ (always). This is true since the tangent function, $f(\theta) = tg(\theta)$, increasing.

From trigonometry, we know equation 34 starts:

$$tan\varepsilon = tan(\alpha_2 - \alpha_1) = \frac{tan\alpha_2 - tan\alpha_1}{1 + tan\alpha_2 * tan\alpha_1} \quad (34)$$

Defining tan(α2) = μ_c and tan(α1) = μ_d, and applying the arctan, we obtain, the equation 35 is addressed:

$$\varepsilon = arctan\left(\frac{\mu_c - \mu_d}{1 + \mu_c * \mu_d}\right) \quad (35)$$

And μ_c being the greater value between μ and μ', and μ_d being the smallest between them. Therefore, 35 is the third main equation of the method, in which ε allows comparing the main basin with the others in terms of shape.

For the main equations, *k* and *j* will have km as their unit, while ε will be considered in degrees. Thereafter, the dependence between the basin shape indicators (*k* and *j*) and the comparison ε between the main basin and each of the others analyzed will be assessed graphically and algebraically. Therefore, we will develop two scatter plots. Since both will have the ε data on the *x*-axis, the first one will have the *k* values on the *y*-axis, while the second will have the *j* data on this same axis. Its trend line will be drawn so that it has the most convenient formation law, that is, the one that most closely approximates the curve resulting from the scattered points.

Figure 6 and 7, with exclusively illustrative purpose, exemplify that situation, using demonstrative values that do not correspond to any actual set of watersheds. Trend lines were obtained in Excel, using the tool designed to obtain approximation functions and considering the function options that best visually fit the sets of scattered points. As for trend lines, we remind that, for a set of *n* points, there will always be a polynomial of degree *n* capable of algebraically describing a curve that passes perfectly through all n points of the set. Preliminary this fact would eliminate the need to use trend lines. However, we must understand that, in Civil Engineering, the most used computational tools have limitations related to the order of approximation functions and their type, making the demand for trend lines not be overcome so easily. The following procedure was used for all points of both graphs during the calculations when applying the method. However, for the deduction of the last necessary formula, it is sufficient to have a point as an example, as shown in Figure 8. In the method's script, the point is horizontally projected onto the trend line so as to reach the region closest to it and keep its image on the axis *y*. Then, another point is created on the curve from the projection so that the tangent line to the curve at the given point is determined. Then, a straight line parallel to the referred one is drawn to pass through the original point, finally calculating the distance between the two straight lines.

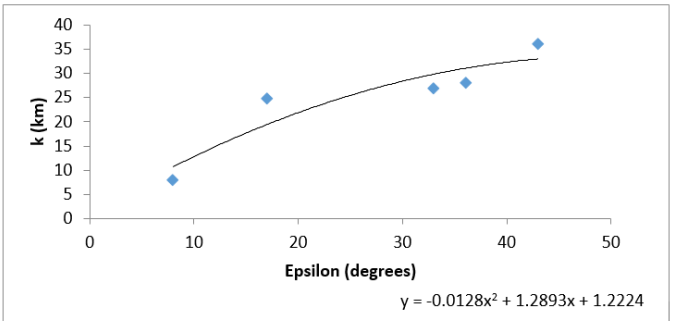

Figure 6 Chart *k x* ε. Source: The authors (2023).

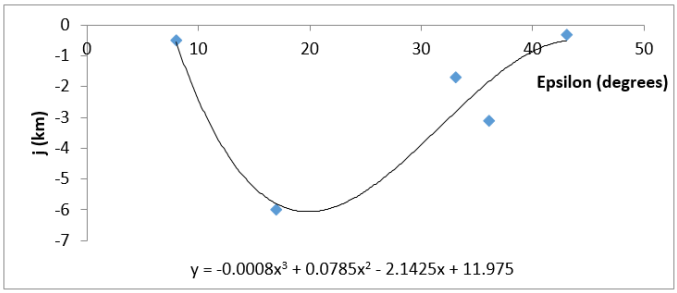

Figure 7 Chart *k x* ε. Source: The authors (2023).

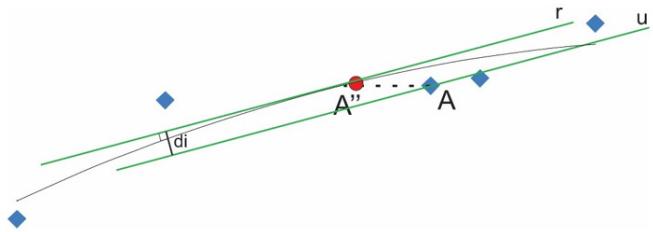

Figure 8 Lines and point projection. Source: The authors (2023).

The point *A* is part of the graph *k x* ε . *A"* and is its horizontal projection on the trend line expressed as a 2nd degree polynomial. *r* is

Citation: Bueno GC, Pereira CE. Calculation method for the classification of water basins shape within a specific basins set. *Int J Hydro.* 2025;9(4):180–187. DOI: 10.15406/ijh.2025.09.00419

the tangent line to point *A"*, and *u* is the line parallel to *r* and passing through the point *A*. Initially, one must define *r*:

Being *f(x)* the function corresponding to the trend line formation law and obtained using computational tools so that this is the function that best approximates the scattered points. Knowing the coordinates of point *A* (the ellipsis that approximates a hydrographic basin) and being *A* (ε ; *k*), *A"* has the same value as *k*, while its value of ε is the one whose application in the equation of the curve results in an image equals *k*. Therefore, if it is point *A"* (ε"; *k*), equation 36 must be completed:

$$f(\varepsilon") = k \quad (36)$$

That is, *k* is the very image of $\varepsilon"$. Thus, to find $\varepsilon"$, just isolate the variable *x* in the trendline equation, putting $y = k$. After calculating said value, you must find the angular coefficient of *r*, which is given through equation 37:

$$m = f'(\varepsilon") \quad (37)$$

Being *m* the angular coefficient of the tangent line to the trend line, and *f'(x)* the first order derivative of *f(x)*.

This means that the angular coefficient of *r* is obtained by substituting $\varepsilon"$ in the variable *x* in the first-order derivative of *f(x)*. As we know that *r* passes through point *A"*, we can calculate the linear coefficient, since we have equation 38:

$$k = m * \varepsilon" + n \quad (38)$$

Where *n* the linear coefficient of the tangent line to the trend line.

Therefore, equation 39 is valid:

$$n = k - m * \varepsilon" \quad (39)$$

Applying equation 37 to equation 39, we obtain equation 40, which represents *r* in $y = m * x + n$.

$$y = f'(\varepsilon") * x + (k - f'(\varepsilon") * \varepsilon") \quad (40)$$

For the line *u*, equation 37 is also valid, since the lines are parallel and their angular coefficients are equal. Since $y = k$ is again at point *A*, through which *u* passes, it is correct to say that what translates into equation 41, which represents *u* in the form $y = m * x + n$, is true.

$$y = f'(\varepsilon") * x + (k - f'(\varepsilon") * \varepsilon) \quad (41)$$

To find the distance *d* between them, consider the geometric formula 42 below:[17]

$$d = \frac{|C - D|}{\sqrt{A^2 + B^2}} \quad (42)$$

Being *A*, *B*, *C*, and *D* the parameters for two straight line equations in the form $A * x + B * y = C$ and $A * x + B * y = D$.

Thus, we have equations 43 and 44 below:

r: $f'(\varepsilon") * x - y = f'(\varepsilon") * \quad - k$ (43)

Where $A = f'(\varepsilon")$, $B = -1$, and $C = f'(\varepsilon") * \varepsilon" - k$.

u: $f'(\varepsilon") * x - y = f'(\varepsilon") * \quad - k$ (44)

Where $A = f'(\varepsilon")$, $B = -1$, and $D = f'(\varepsilon") * \varepsilon - k$.

So, equation 45 is true:

$$d_i = \frac{\mu_l - \mu_m}{\sqrt{(f'(\varepsilon"))^2 + 1}} \quad (45)$$

Being d_i the observed graphic distance for each basin, μ_l the highest value between $f'(\varepsilon") * \varepsilon" - k$ and $f'(\varepsilon") * \varepsilon - k$, and μm the smallest between them.

Similarly, this holds for *j*, where μ_l is the largest value between $f'(\varepsilon") * \varepsilon" - j$ and $f'(\varepsilon") * \varepsilon - j$, and μm the smallest between them.

After that, we can calculate the arithmetic mean of the distances using equation 46 (BUSSAB & MORETTIN, 2013):

$$\overline{d_i} = \frac{\sum_{i=1}^{N} d_i}{N} \quad (46)$$

Being $\overline{d_i}$ the average of the distances, and *N* is the number of basins studied.

Finally, d_i and $\overline{d_i}$ can be compared in the following sense: $d_i < \overline{d_i}$ (basin inside the mean), $d_i > \overline{d_i}$ (basin outside the mean) and $d_i = \overline{d_i}$: basin in the mean.

Some practical remarks about this script are: for the main basin, $\varepsilon = 0$ (always), since the corresponding point will be compared with itself; the study should not evaluate less than three basins, since the method will lose its effect as the trend line would be a straight line passing precisely through the single point or the two points in dispersion, so that the distances measured between the two parallel lines in the graphs would be null.

Deduction of the vertical projection equations

Similarly to the development of the horizontal projection of points, the vertical projection can be introduced. Therefore, Figure 9 demonstrates this mathematical tool's geometric and analytical principle within differential and integral calculus. In the figure below, *A* is the original point, *A'* is the projected point, *s* is the line tangent to the curve at *A'*, and *v* is the line parallel to *s* and passing through *A*.

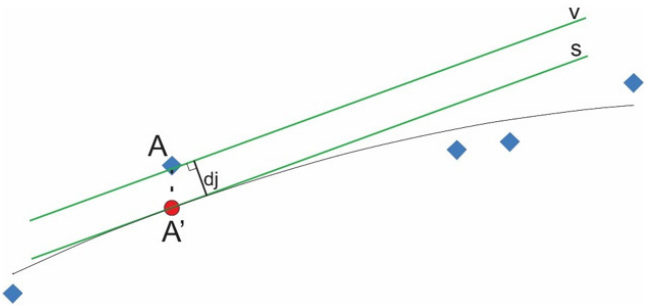

Figure 9 Vertical projection of the point and auxiliary lines. Source: The authors (2023).

Thus, under the same line of reasoning as the deduction for the projection on *y*, the projection on *x* can be considered to obtain equations 47 and 48, which refer to *k*, but which equally serve to *j*:

$$k = f'(\varepsilon) * \varepsilon + n_1 \quad (47)$$

$$k' = f'(\varepsilon) * \varepsilon + n_2 \quad (48)$$

Citation: Bueno GC, Pereira CE. Calculation method for the classification of water basins shape within a specific basins set. *Int J Hydro.* 2025;9(4):180–187. DOI: 10.15406/ijh.2025.09.00419

Being k the image of the original point, k' the image at the projected point, ε the value on the axis x for both points, n_1 the linear coefficient of the line passing through the original point, and n_2 the linear coefficient of the tangent line to the trend line.

Therefore, we proceed to equations 49 and 50 for the real and projected points.

$$n_1 = k - f'(\varepsilon) * \varepsilon \quad (49)$$

$$n_2 = k' - f'(\varepsilon) * \varepsilon \quad (50)$$

In this case, we obtain equations 51 and 52:

$$f'(\varepsilon) * \varepsilon - k = f'(\varepsilon) * \varepsilon - k \quad (51)$$

$$f'(\varepsilon) * \varepsilon - k' = f'(\varepsilon) * \varepsilon - k' \quad (52)$$

Being $A = f'(\varepsilon)$, $B = -1$, $C = f'(\varepsilon) * \varepsilon - k$, and $D = f'(\varepsilon) * \varepsilon - k'$, we can conclude that equation 53 is valid, which calculates the distance d_j between lines for the projection on x:

$$d_j = \frac{\mu_n - \mu_o}{\sqrt{(f'(\varepsilon))^2 + 1}} \quad (53)$$

Being μ_n being the largest value between $f'(\varepsilon) * \varepsilon - k$ and $f'(\varepsilon) * \varepsilon - k'$, and μ_o being the smallest value.

Deduction of the compound projection equation

After specific deductions of d_i and d_j, we can find the distance d_c given by the composite projection, which is the geometric mean between the distances calculated for the horizontal and vertical projections. Equation 54, in turn, demonstrates how this geometric mean is calculated:[18]

$$d_c = \sqrt{d_i * d_j} \quad (54)$$

Thus, we can compare individual and average distances for the three types of calculated distances: the horizontal projection (d_i), the vertical projection (d_j) and the composite projection (d_c).

Calculating the semi-axes of the ellipsis

We obtain measurements as follows: the distance between the watershed mouth and the farthest point from it is measured as the longest axis of each approximation ellipsis. The parameter a, in turn, is half of the distance. Parameter b is determined as the arithmetic mean between the distances (right and left) from the central point of the major axis to the edges of the respective basin, such that the measurements are made on the axis perpendicular to the major axis. Such segments are highlighted in Figure 1 in black. We used the AutoCAD software in the measurements on a contour map.

Results

Table 1 summarizes the data from the measurements made in the three hydrographic basins studied. These measurements allowed the calculation of the semi-axes of the approximation ellipses and their use as input parameters for the script calculations.

Table 1 Dimensions of watersheds. Source: The authors (2023)

Major axis (2*a) (km)	Longest Perpendicular Distance (km)	Smallest perpendicular distance (km)
Basin 1 87,5658	19,0106	9,7155
Basin 2 153,9927	50,1284	4,8250
Basin 3 153,9927	17,5906	12,9690

Table 2, on the other hand, provides the calculated values of the distances resulting from the horizontal projection for basins 1, 2, and 3, representing parameters k and j, according to the indexes indicated.

Table 2 d_i for k and j (lu). Source: The authors (2023)

d_{i1k}	0,00023866
d_{i2k}	0,00000546
d_{i3k}	0,00002206
d_{i1j}	0,00027826
d_{i2j}	0,00003780
d_{i3j}	0,00091246

Table 3, in turn, provides the average values of the distances resulting from the horizontal projection concerning parameters k and j.

Table 3 $\overline{d_i}$ to k and j (lu). Source: The authors (2023)

$\overline{d_{ik}}$	0,00008873
$\overline{d_{ij}}$	0,00040951

Concerning these results, we draw attention to what is written in Table 4, that is, basin 1 is considered outside the mean for k. Basin 1 is inconsistent with the average behavior, represented by the trend line. At the same time, it can be seen that for j this is true for basin 3.

Table 4 Analysis of the horizontal projection results. Source: The authors (2023)

As for k x ε		
Basin 1	OUTSIDE THE MEAN	NOT IN THE MEAN
Basin 2	INSIDE THE MEAN	NOT IN THE MEAN
Basin 3	INSIDE THE MEAN	NOT IN THE MEAN
As for j x ε		
Basin 1	INSIDE THE MEAN	NOT IN THE MEAN
Basin 2	INSIDE THE MEAN	NOT IN THE MEAN
Basin 3	OUTSIDE THE MEAN	NOT IN THE MEAN

As for the vertical and composite projections calculations, the procedures were only performed for parameter k, to save time. However, the same can be applied for j. All these results have been tabulated, starting from Table 5, which is designated to introduce the values of calculated values of the distances resulting from the vertical projection for basins 1, 2, and 3.

Table 5 d_j to k (lu). Source: The authors (2023)

d_{j1k}	0,00003251
d_{j2k}	0,00000548
d_{j3k}	0,00003522

To keep the reasoning behind the work, Table 6 presents the *status* of "outside the mean" for basins 1 and 3, meanwhile basin 2 is classified as "inside the mean". In this case, basin 3 is the only basin with a different configuration from that to be found in Table 4.

Table 6 Analysis of the vertical projection results. Source: The authors (2023)

As for k x ε		
Basin 1	OUTSIDE THE MEAN	NOT IN THE MEAN
Basin 2	INSIDE THE MEAND	NOT IN THE MEAN
Basin 3	OUTSIDE THE MEAN	NOT IN THE MEAN

Citation: Bueno GC, Pereira CE. Calculation method for the classification of water basins shape within a specific basins set. *Int J Hydro.* 2025;9(4):180–187. DOI: 10.15406/ijh.2025.09.00419

For the compound projection, Table 7 has been designed to introduce the calculated distances for parameter *k*, including the same 3 basins.

Table 7 d_c to *k* (lu). Source: The authors (2023)

d_{c1k}	0,00008808
d_{c2k}	0,00000547
d_{c3k}	0,00002787

To conclude the results section, Table 8 shows basins 1, 2 and 3 in the same *status* configuration that can be found in Table 4 for parameter *k*.

Table 8 Analysis of the composite projection results. Source: The authors (2023)

As for *k* x ε		
Basin 1	OUTSIDE THE MEAN	NOT IN THE MEAN
Basin 2	INSIDE THE MEAN	NOT IN THE MEAN
Basin 3	INSIDE THE MEAN	NOT IN THE MEAN

In this regard, while basins 1 and 2 maintain their respective status of outside and inside the mean in all projections, basin 3 shows changes, going from inside to outside the mean from the horizontal to the vertical projection, and returning to inside the mean in the composite projection.

Discussion

As Table 1 stablishes, the descending order for the major axis, the longest and the smallest perpendicular distance is, respectively: basin 2, basin 1 and basin 3; basin 2, basin 1 and basin 3; basin 3, basin 1 and basin 2. Furthermore, basin 3 is the one in which the longest perpendicular distance is the closest to the smallest, meanwhile basin 2 is the one with the biggest disproportion related to the perpendicular distances. That information is relevant to demonstrate differences in the basins' shapes are significant, especially when basins 2 and 3 are placed in a pair to be directly compared.

As shown by Table 2, the descending order for d_i related to the *k* x ε analysis is basin 1, basin 3 and basin 2, revealing basin 1 has the furthest position to the trend line in terms of the horizontal projection, and that basin 2 has the closest. Meanwhile, the descending order for d_i in terms of the *j* x ε analysis is basin 3, basin 1 and basin 2, which indicates that basin 3 is the one with the furthest position to the trend line in terms of the horizontal projection, and that basin 2 has the closest position. So, the fact that analyses based on different parameters can bring notably different results is demonstrated by that situation and should be emphasized. Those results are useful to identify the differences in the behavior of curves stablished for *k* x ε and *j* x ε, which is something predictable, since the trend lines are oriented by fundamentally different parameters.

Table 3 shows that the average value of the d_i for the *j* x ε analysis is higher than the one for the *k* x ε analysis. That is a predictable result, due to what is demonstrated in Table 2, in which bigger values of the distance can be seen for basins 2 and 3, since basin 1 is an exception, as it keeps its values of d_i reasonably close for both *k* x ε and *j* x ε analyses. Those data are important to point out the difference in susceptibility of basins' geometrical conditions to the change of parameters that guide the analysis. It is not something that happens particularly in this case. It is something that is shown as possible to happen in many other analyses.

In addition to stablishing there are no values of d_i that are in in the mean, or exactly equal to the average values calculated for each one of the analysis, Table 4 reinforces how different choices of the parameter to substantiate the analyses can bring different results. Objectively, basins 1 and 3 switch conditions from *k* x ε to *j* x ε analysis. Basin 1 holds an "outside the mean" state in terms of *k* x ε and an "inside the mean" in terms of *j* x ε. Meanwhile, the opposite happens to basin 3, and basin 2 keeps its state stable as "inside the mean". That shows how the difference in sensitivity of basin's condition to the change of the parameter to guide the analysis, as described in the previous paragraph, has an impact on the *status* comparison between *k* x ε and *j* x ε analyses, since the average distance is conditioned by all the distances calculated, which are directly influenced by that sensitivity.

At this point, there are some comparisons that need to be made. The first one includes Table 2, 5 and 7. In descending orders, as d_c is the geometric mean of d_i and d_j, it is always going to exist between d_i and d_j. So that is why all the results from Table 7 occupy the second position in the orders to be stablished. To conclude this reasoning, for basin 1, d_i is the biggest distance calculated, meanwhile d_c is the second one, and d_j is the smallest. For basin 2, d_j is the biggest distance calculated, meanwhile d_c is the second one, and d_i is the smallest. For basin 3, the same configuration that appears in the case of basin 2 is verified to be true. This is a relevant result, since it means that, in terms of *k* x ε analysis, basin 1 presents greater proximity to the curve when it comes to the vertical projection (d_j is the smallest distance), meanwhile basins 2 and 3 reveal that they have that condition in relation to horizontal projection (d_i is the smallest distance).

One point to be emphasized is that the difference between d_i and d_j for basin 2 is very small, at least if compared to what happens to basins 1 and 3. On the other hand, basin 1 suffers the largest variation between both distances. Those data reiterate that the behavior of each basin when compared to the behavior predicted by the trend line changes due to the chosen projection. In relation to the basins this work focuses on, that information is relevant to demonstrate greater regularity of basin 2 in terms of its conditions when compared to the trend line's behavior, since d_i and d_j present more remarkable proximity for that basin. Simultaneously, basin 1 occupies the last position in the aspect of regularity, which leads basin 3 to automatically get the second.

Furthermore, the comparison between Table 4, 6 and 8 reveals that the choice of the projection also has an influence on the result supposed to classify basins as "inside the mean", "in the mean" or "outside the mean". Basin 1 keeps its state as "outside the mean" in every case of projection. Basin 2 keeps its behavior as an "inside the mean" in every case. However, basin 3 presents an "inside the mean" condition for the horizontal and the composite projections (d_i and d_c, respectively) and an "outside the mean" one for the vertical projection (d_j). The significance of those data is to emphasize that, besides presenting the greatest regularity in relation to d_i and d_j, basin 2 has is perfectly regular in the *status* classification, being stablished as "inside the mean" to any chosen projection (in the *k* x ε analysis). Meanwhile, basin 1, the basin with the greatest irregularity in d_i and d_j comparison, presents regular "outside the mean" *status* for all the projections (for *k* x ε analysis). That leads basin 3 to be the only one to revel irregularity in its *status*.[19]

Conclusion

Once the calculations with the equation have been completed, the method is applicable effectively for a set of watersheds, as proposed (minimum of 3 watersheds). Moreover, the equations for the trend lines should be adapted to reproduce accurately average behavior of the set of ellipses, thus minimizing the errors inherent to the calculation process. Finally, it is also interesting to note the

possible need to subdivide regions with many river basins into smaller sets, avoiding failures during the graphical analysis. Moreover, the different results each analysis can offer by being developed based on a different parameter and using distinct projections reveal how applying a variety of methods can make the work of classifying basins from a certain group more robust. For that, it is important to elaborate a comparison between the different results obtained. In the aspect of the 3 basins, general diagnosis for k x ε analysis stablishes that basin 1 presents the most divergent geometrical behavior in relation to trend line curve, since it is classified as "outside the mean" for all 3 projections. Under the same analysis, basin 2 is the one with the closest geometrical behavior in terms of proximity to the trend line curve. Basin 3, on the other hand, stands out as the most inconstant one. For the briefer j x ε analysis, the overall classification changes, revealing basins 1 and 2 as "inside the mean" and basin 3 as "outside the mean" for horizontal projection.

Acknowledgments

None.

Conflicts of interest

The author declares there is no conflict of interest.

References

1. Carbonari FI. *A gestão dos recursos hídricos* (Tese de especialização). Faculdade de Engenharia Mecânica, Universidade Estadual de Campinas, Campinas; 1997.
2. Santos CL, Silva OG, Vital SRO, et al. Análise da suscetibilidade à ocorrência de enchentes e inundações na bacia do Rio Jaguaribe - João Pessoa/PB. *Revista Brasileira de Geografia Física*. 2018;11(5):1876–1888.
3. Soares GAS, Galvíncio JD. Uso do LiDAR para avaliar os padrões hídricos de bacias em áreas urbanas: caracterização fisiográfica da bacia do Rio Beberibe - PE. *Revista Brasileira de Geografia Física*. 2020;13(7):3659–3674.
4. Alves MEP, Fan FM, Paiva RCD, et al. Assessing the capacity of large-scale hydrologic-hydrodynamic models for mapping flood hazard in southern Brazil. *Revista Brasileira de Recursos Hídricos*. 2022;(27)(9).
5. Lima BM, Lima LC. Forma da bacia. In: PJO. Machado editor. *Diagnóstico físico-ambiental da bacia hidrográfica do córrego São Pedro: um exercício acadêmico de gestão dos recursos hídricos*. Ubá: Geographica, 2010. p. 69–72.
6. Schwaab M, Pinto JC. *Análise de dados experimentais I: fundamentos de estatística e estimação de parâmetros*. v. 1. Rio de Janeiro: E-Papers, 2007.
7. Costa Neto PLO. *Estatística*. 2 ed. São Paulo: Edgard Blücher, 2002.
8. Santos LHO. *As paisagens das terras altas da Mantiqueira: uma perspectiva geográfica*. São Paulo: Haikai, 2020.
9. Ab'Sáber AN. *Os domínios de natureza no Brasil*. 7 Ed. São Paulo: Ateliê; 2003.
10. Sant'Anna Neto, J. L.. Decálogo da climatologia do sudeste brasileiro. *Revista Brasileira de Climatologia*. 2005;1(1):43–60.
11. Instituto Nacional de Meteorologia – INMET, 2019?.
12. Sistema Estadual do Meio Ambiente e Recursos Hídricos – SISEMA, 2019.
13. Instituto Brasileiro de Geografia e Estatística – IBGE, 2019.
14. Winterle P. *Vetores e geometria analítica*. 2 ed. São Paulo: Pearson Education do Brasil, 2014.
15. Nascimento SV. *Desvendando os segredos do triângulo retângulo e descobrindo curiosidades até hoje não conhecidas*. 1 ed. Rio de Janeiro: Gramma, 2018.
16. Quaranta F, Lourenço EG, Alves A, et al. Geometria e aritmética combinam com equações do 2º grau?. *Holos*. 2013;(6)(29):149–160.
17. Bourchtein A, Bourchtein L, Nunes GS. *Geometria analítica no plano*. São Paulo: Edgard Blücher; 2019.
18. Silva ALC. *Introdução à análise de dados*. 2 ed. Rio de Janeiro: E-Papers, 2011.
19. Bussab WO, Morettin PA. *Estatística básica*. 8 ed. São Paulo: Saraiva; 2013.