Journal of Physics: Conference Series **1448** (2020) 012021 doi:10.1088/1742-6596/1448/1/012021

Regular time series for physical-chemical water parameters: Total suspended solids and hydrogen potential

J F Ortiz[1], M J Mesa[1], and O E Molina[1]
[1] Universidad del Quindío, Armenia, Colombia

E-mail: omolina@uniquindio.edu.co

Abstract. The Quindío river is the main water source of the Quindío department, Colombia, which supplies water to 410000 people. Therefore, the study of its water quality is extremely important. If pollution is not monitored and controlled in this water source, diseases could affect the population of people that use the river as a source of water. For this reason, fieldwork was carried out in an area of the Quindío river where the total suspended solid and hydrogen potential parameters were studied. In order to analyze the data that was obtained in the fieldwork, the spine interpolation needed to be performed to generate a new sample where the spatial data was equidistant. To determine the relationship between the total suspended solid and hydrogen potential parameters, two regular time series were built and the second-order vector autoregressive model was adjusted. The model order was suggested by the statistical package Eviews.

1. Introduction

Colombia has two large water sources which are the Pacific and the Atlantic Ocean. It also has 737000 surface bodies of water (rivers, lakes, wetlands, and lagoons) and its water supply exceeds 2000 km^3 per year, which makes Colombia a global water power [1]. Looking at the department of Quindío, it has approximately 140 creeks and 11 rivers. The main source of water is The Quindío river that originates in the mountains east of Salento, and flows in a southwest direction at the confluence with Barragán river giving rise to the La Vieja river.

Water is an indispensable component for all biological processes and it is also the component that gives rise to life on the planet. In addition, the hydrological cycle of water maintains the climate and atmospheric stability of planet earth. The main function of water is the transportation of nutrients that allows for the development of human life, animal life, as well as the process of photosynthesis in plants, and all life forms here on earth [2,3]. As time goes on, water loses its purity and quality due to industrial activities, population growth and unnecessary water expenditure. In this way, it is absolutely important to preserve the quality of all water sources in order to conserve life in all ecosystems.

To measure water quality, different physical-chemical parameters such as dissolved oxygen (DO), hydrogen potential (pH), chemical oxygen demand, electrical conductivity, total suspended solids, among others [4] are taken into account. In Colombia there are different indexes that measure water quality. Among them are the water quality index (WQI) and the water quality risk index (WQRI). The first one refers to a single value (between 0 and 1) that indicates water quality by assessing physical, chemical and microbiological conditions. While the second one establishes a percentage in which 0% to 5% water is suitable for human consumption. Water between 70.1% and 100% becomes unviable for health. There are also laws and rules governing everything related to water sources, such as resolution

0631 (a resolution from 2015) in which the maximum permissible values for occasional water sheds were presented [5].

In this research, an analysis of the total suspended solid (TSS) and pH parameters were evaluated and the data was taken from The Quindío river in the sector La María. The data was gathered during 8 weeks in the rainy season and 21 sampling stations were taking into account. For pH, the results of the stations were averaged and the polynomial of these new points was obtained by means of spline interpolation. Similarly, for the TSS, the median data of each station was used to obtain a second polynomial (because in the last sampling campaigns there were heavy rains in the department of Quindío in the months of November and December 2018, TSS took very high values and for that reason, scientists decided to work with the median). Finally, the length of the study area was distributed in equal parts with the aim of taking 50 data points distributed regularly in space which generated a regular time series.

2. Materials and methods

The study area extends along approximately 8.5 kilometers and the first station is located 500 meters above the bridge of La María. Most of the monitoring parts are concentrated in this part of the river because in the first two kilometers there is a considerable source of pollution that is caused by human settlements, tanning companies and commercial processes. Figure 1, shows the study area, indicating the monitoring points (points 19 and 20 do not appear in the image).

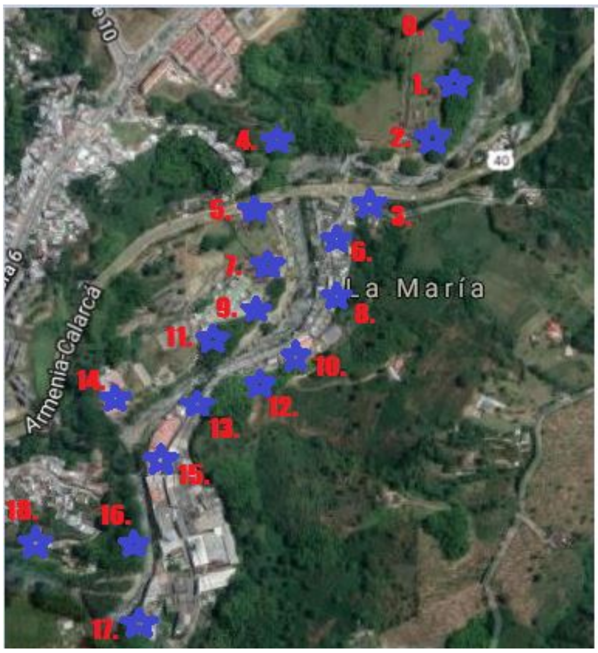


Figure 1. Study area Calarcá, Armenia, Colombia.

The physical-chemical parameters of the water that were studied TSS and pH, since these parameters are taken into account by the Colombian regulations in different scenarios, both for water modeling [6], as well as to determine water quality. The pH influences aquatic ecosystems and industrial processes. For this reason, it is important to keep their values within acceptable ranges. When TSS reaches very high values, it causes damage to animal and plant life. Also, the TSS can stay inside of the water, which leads to its accumulation and subsequent deterioration of oxygen [7].

According to the Colombian technical standard NTC-ISO 5667-1 [8] the sampling technique to be used in this study is the point sampling technique, since each sample is normally a representation of the water quality only at the time and place where it is taken. This type of sampling is recommended in research that seeks to detect possible contamination, the scope of the contamination and to determine whether or not the water quality meets the limits related to average quality [8].

3. Interpolation of the data

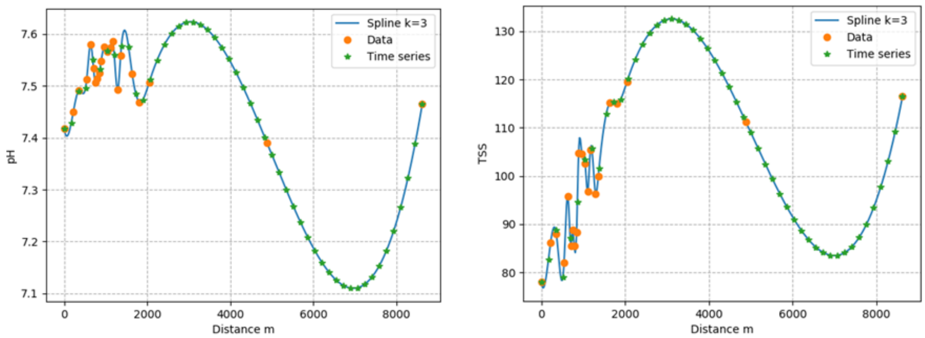
The spline interpolation technique is used to obtain a fragmented function with third-degree polynomials for each interval between the nodes. These third-degree polynomials Equation (1) have the following form for each i ϵ I:

$$f_i(x) = a_i x^3 + b_i x^2 + c_i x + d_i \tag{1}$$

This technique allows researchers to smoothly and continuously join a series of points, avoiding the possible oscillations that would occur with high-grade polynomials [9].

3.1. Applying spline interpolation to sampling data

The results obtained from the monitoring were used to perform cubic spline interpolation and to have a better representation of the 8 pH studies. The data was averaged at each station and the interpolating polynomial was obtained for the new data. Similarly, the same method was used for the TSS parameter, with the difference that the median was used because there were a few weeks where outlier data was presented due to heavy rains [9,10]. Finally, with the two interpolant polynomials and the distances between the monitoring points were taken into account, a total of approximately 8.5 kilometers. A partition of the distances was made with the aim of taking 50 regular data points and building a series regular time sections for each parameter Figure 2(a) and Figure 2(b).



(a) (b)

Figure 2. Interpolant polynomials based on distances and divided regularly into 50 sections. Interpolator polynomial for (a) the average pH and (b) median TSS.

3.2. Autoregressive model

To understand the interrelationship of the measured variables, TSS and pH mentioned above, two regular time series were generated and the second-order vector self-regressive model was adjusted [11]. Initially the graphical stationarity test is done, indicating that the inverse roots of the characteristic equation are within the unit circle. Therefore, it is concluded that the series are stationary as shown in Figure 3. In addition, analytically, the inverses of each of the roots is less than 1 in module, as shown Table 1.

As mentioned earlier, the stationarity of the vector series is checked. As the series are regular and stationary, then it is possible to adjust a vector self-regressive model, in this case of order 2, as suggested by the statistical package. The model is described by Equation (2) and Equation (3).

$$pH_t = -0.050173502\ pH_{t-1} + \varepsilon_{1,t} \tag{2}$$

$$SST_t = 1.556550995\ pH_{t-1} + 10.84124632\ pH_{t-2} + \varepsilon_{2,t}, \tag{3}$$

where pHt-1 and pHt-2 refer to the previous stage and the previous two stages, respectively. $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ refer to two random errors. This model interprets the following: the pH depends only on the pH at the previous stage or station in a negative way, plus a random error and the TSS depends on the pH in the previous two stages positively plus another random error.

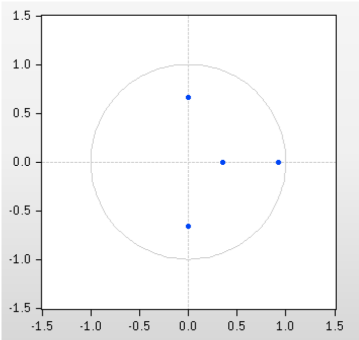


Figure 3. TSS and pH series stationarity test in the Eviews statistical program.

Table 1. Inverse of the roots.

Root	Modulus
0.926227	0.926227
0.004582-0.664112i	0.664128
0.004582+0.64112i	0.664128
0.350334	0.350334

No root lies outside the unit circle. VAR satisfies the stability condition.

The above model describes the interdependence of the pH and TSS variables. Note that not all coefficients of the model have been taken, i.e. the pH would be expected to depend on itself and TSS in the previous two stages or stations. These coefficients are chosen with a probability of 90% confidence, for this, the t-statistic module must be less than 1.7.

Table 2 and Table 3 show the coefficients and values of the t-statistics. In addition, the R-squares of the series are presented in Table 4.

Table 2. Coefficients.

Serie	pH_{t-1}	pH_{t-2}	SST_{t-1}	SST_{t-2}
pH	-0.050173502	-0.459013330	-0.003008143	0.002912255
SST	1.556550995	10.84124632	1.335899034	-0.380584769

Table 3. t- statistic.

Serie	pH_{t-1}	pH_{t-2}	SST_{t-1}	SST_{t-2}
pH	-0.38425	-3.48795	-1.85247	1.83830
SST	0.13629	0.94184	9.40538	-2.74656

Table 4. R-square.

Serie	R- square
pH	0.2607
SST	0.9532

Since the R-square of the TSS variable is 0.9532, the model explains at 95.32% that variable. In addition, it can be concluded that the model explains the pH variable by 26.07%, so the estimates are not as reliable for such a variable. On the other hand, it is concluded that the explanatory or prediction variable is the pH variable, since, it is the one that affects the TSS variable and for the same reason it is said that the TSS variable is the explained variable [12,13].

4. Conclusions

Spline interpolation allowed new measurements to be generated in the Quindío river by using the pH and TSS parameters in different places. These methods were fundamental in generating the two regular time series. Thus, the measurements are more accurate, because the spline doesn't have curvature changes as large as those presented with other interpolation techniques. In addition, spline interpolation

Journal of Physics: Conference Series **1448** (2020) 012021 doi:10.1088/1742-6596/1448/1/012021

has the advantage of allowing you to switch between distinctly continuous and continuous polynomial domain sections in a way that improves accuracy.

It is important to clarify that the results obtained in this work are only conclusive for the variables studied and type of sampling used. In addition, vector autoregressive models are a data-driven tool used to show the interrelationship of analyzed stationary series whose observations are regular. These models are made up of a system of equations and differences that de-scribe the relationships between variables, but do not indicate the causality between the variables.

References

[1] Pineda C and Samper M 2008 *Agua y vida en Colombia* (Colombia: Editorial Politécnico Grancolombiano)
[2] Aguilar A 2004 *Geografía general* (México: Editorial Pearson Educación) p 113
[3] Vaquero M 2012 *Aguas para la salud: pasado, presente y futuro* (Madrid: Consejo superior de Investigaciones Científicas)
[4] Ramírez C A *2016 Calidad del agua. Evaluación y diagnóstico* (Colombia: Sello Editorial de la universidad de Medellín) p 149
[5] Ministerio de Ambiente y Desarrollo Sostenible 2015 *Resolución N° 0631. Por la cual se establecen los parámetros y los valores límites máximos permisibles en los vertimientos puntuales a cuerpos de aguas superficiales y a los sistemas de alcantarillado público y se dictan otras disposiciones* (Colombia: Ministerio de Ambiente y Desarrollo Sostenible)
[6] Ministerio de Ambiente y Desarrollo Sostenible 2015 *Decreto Ley N° 3930, 2010. Por el cual se reglamenta parcialmente el Título I de la Ley 9ª de 1979, así como el Capítulo II del Título VI -Parte III- Libro II del Decreto Ley 2811 de 1974 en cuanto a usos del agua y residuos líquidos y se dictan otras disposiciones* (Colombia: Ministerio de Ambiente y Desarrollo Sostenible)
[7] Ruiz D C 2017 *Método de estimación de sólidos suspendidos totales como indicador de la calidad del agua mediante imágenes satelitales* (Bogotá: Universidad Nacional de Colombia)
[8] Instituto Colombiano de Normas Técnicas y Certificación (ICONTEC) 2013 *Muestreo. Directices para el diseño de programas de muestreo, Norma Técnica Colombiana, NTC-ISO 5667-1* (Colombia: Instituto Colombiano de Normas Técnicas y Certificación)
[9] Trincado G and Vidal J 1999 Aplicación de interpolación “spline” cúbica en la estimación de volumen *Bosque* **20(2)** 3
[10] Dominguez F and Nieves H 2014 *Métodos numéricos aplicados a la ingeniería* (México: Grupo Editorial Patria) p 370
[11] Box G, Jenkins G, Reinsel G and Ljung G 2016 *Time series analysis* (New Jersey: Jhon Wiley & Sons)
[12] Reinsel G 1993 *Elements of multivariate time series analysis* (New York: Springer-Veelag)
[13] Wei W 1990 *Time series analysis univariate and multivariate methods* (United State of America: Person Adisson Wesley)