TEKNOS Revista Científica

Water quality assessment in barbacoas bay - Cartagena using multivariate statistical methods

Evaluación de la calidad de agua en la bahía de barbacoas Cartagena utilizando métodos estadísticos multivariables

Recibido: 15-feb-17 – **Revisado:** 30-abr-17
Aceptado: 30-may-17 – **Publicado:** 30-jul-17

Ildefonso Baldiris-Navarro
Researcher, Fundación Universitaria Tecnológico Comfenalco, Sede A, Barrio España Cr 44 D Nº 30A – 91, Cartagena, Colombia.
ibaldiris@tecnocomfenalco.edu.co

Jorge Sánchez Aponte
Researcher, Sena, Centro Internacional Náutico Fluvial y Portuario, Mamonal, km 5, Cartagena, Colombia.
jhsancheza@sena.edu.co

Daylin Cabeza
Industrial Engineering Student, Fundación Universitaria Tecnológico Comfenalco, sede A, Barrio España Cr 44 D Nº 30A – 91, Cartagena, Colombia.
dcabezat@tecnocomfenalco.edu.co

Juan Carlos Acosta
Researcher, Fundación Universitaria Tecnológico Comfenalco, sede A, Barrio España Cr 44 D Nº 30A – 91, Cartagena, Colombia.
jacosta@tecnologicocomfenalco.edu.co

Abstract: In this study, statistical multivariable technics were applied to an eleven years' complex data base, in order to extract important information about the water quality in a polluted bay in the city of Cartagena, called Barbacoas. This bay receives hundreds of tons of sediments coming from Canal del Dique, a part of the Magdalena river in Colombia. Kruskal Wallis test and principal components analysis were used to find similarities, dissimilarities and Compliance with Colombian law among the different monitoring points. It was found that phosphorus, salinity and solids are out of the range set by the Colombian law. Principal components analysis showed that eight variables explain 64% and 69 % of the variability of data for dry and rainy season.

Key words: Barbacoas bay, Canal del dique, Principal components analysis, Kruskal-wallis.

Resumen: En este estudio, fueron aplicadas técnicas de estadística multivariada a una base de datos de 11 años, para extraer información importante sobre la calidad del agua en una bahía contaminada de la ciudad de Cartagena, llamada barbacoas. Esta bahía recibe cientos de toneladas de sedimentos provenientes del canal del dique, que es una rama del rio magdalena en Colombia. El test de kruskal wallis y el análisis de componentes principales se usaron para encontrar similitudes, diferencias y cumplimiento de la ley colombiana entre los diferentes puntos de muestreo. Se encontró que el fosforo, la salinidad y los sólidos están por fuera del rango de aceptación establecido por la ley colombiana. El análisis de componentes principales mostro que ocho variables explican el 64% y 69%de la variabilidad de los datos para las épocas secas y de lluvias respectivamente.

Palabras clave: bahía de Barbacoas, Canal del dique, análisis de componentes principales, Kruskal-wallis.

1. INTRODUCTION

Since 1958 Barbacoas bay is receiving hundreds of tons of sediments coming from two mouths of the Canal del Dique called Matunilla and Lequerica. The Canal del Dique is a man-made canal that communicates the coast with the inner part of Colombia. Because of the mix of marine water and fresh water the bay has become an estuary. The aim of opening these mouths in The Canal was to decrease the amount of sediments arriving to Cartagena bay, one of the most important harbor in Colombia (Giraldo et al., 2009).

Near Barbacoas bay is located the Rosario islands natural park which is a protected area due to its ecosystemic functions, essential ecological processes, it also acts as a coral barrier that mitigates the impact of coastal erosion, is the habitat of fishes and invertebrates of commercial value and besides has beautiful and attractive landscapes based on the beauty of coral reef that encourages ecotourism. The water discharge flow through the Canal del Dique has been calculated in about $397 m^3 s^{-1}$, of which approximately 35% is discharged into Barbacoas Bay through Matunilla and Lequerica canals, and 17.5% is discharged through an outlet canal further south of this bay, which is called Correa.

Near the discharges of the river there are several shrimp industries and communities directly associated with sediments loading at the mouths of the river. Due to this location, near to the Rosario Islands, the discharge into Barbacoas Bay and that through Correa outlet canal are of particular concern.(Restrepo, Park, Aquino, & Latrubesse, 2016)(Moreno-Madriñán et al., 2015)(Pinilla, Gutiérrez, & Ulloa, 2007).

The simplest evaluation of water quality might be done using basic statistics (mean, SD, etc.) and some graphical aids. However, the problem increases with the number of data because information may be missed; then the use of multivariate techniques and data reduction are almost mandatory to achieve satisfactory results. The use of Principal components analysis (PCA) and others techniques to water quality assessment has increased in the last years, mainly due to the need to obtain appreciable information from data for decision making.(Alberto et al., 2001)(Chow et al., 2016). The aim of this study is to analyze physico-chemical and microbiological parameters of water quality from Barbacoas bay to evaluate information about the similarities, dissimilarities and compliance with Colombian law among the different monitoring stations and to know the impact of the pollution sources on the water quality parameters.

2. MATERIALS AND METHODS

2.1. Study area

Barbacoas bay is located in the Colombian Caribbean at 75°31' W and 75°43 W longitude and latitude 10°07' N and 10°15' N. The bay has an area of 120 km^2 and it is influenced by the discharges of Canal del Dique, which is a part of Magdalena river in Colombia. The three sampling points (Matunilla, Lequerica and Baru) were designed by the Corporacion autonoma del canal del dique, which is the local environmental authority.

The Canal transports a great number of solids, coliforms and nutrients coming from village and towns along the way of the river. In the zone, there are also big shrimp companies and agricultural activities where the use of fertilizers and other chemicals is a common practice that may cause deterioration in the bay's water quality and consequently in coral habitat(Acosta & Baldiris, 2015).

Fig. 1. Location of water quality sampling stations.

2.2. Sample collection

The sampling stations in Barbacoa bay are depicted in Fig. 1. Two of the three points are near to Canal del Dique discharges (Matunilla and Lequerica), the other one (Baru) is near to a natural park that is visited by hundreds of tourists every day. The samples were taken twice a year according to the season (dry-rainy), APHA protocols were followed in order to have accurated results. (Chow et al., 2016)(Bhuiyan, Rakib, Dampare, Ganyaglo, & Suzuki, 2011).

2.3. Analytical methods

The selected water quality parameters were dissolved oxygen (DO), Total coliform (TC), pH, water temperature (T), 5-day biochemical oxygen demand (BOD5), nitrate (NO3-), total phosphorus (P), total solids (TS), Salinity(Sal) and ammonium(NH4+).

The temperature, pH, Salinity and DO concentrations were measured on-site by a thermometer, pH, salt and DO meter (using HACH sension 5465011 portable), respectively. The BOD5 was determined by Winkler method and TS were determined gravimetrically at 105-110 °C. NO3- and P were analyzed by cadmium reduction and ascorbic acid method (using Varian Cary 100 UV -vis spectrophotometer), respectively. The membrane filter technique was used to determine the number of colony forming units per 100 mL (cfu/100 mL) of total coliforms in sampling water(American Public Health Association, American Water Works Association, & Water Environment Federation, 1999). Table 1 has a summary of the standard methods used in this investigation.

Table 1. Summary of used standard methods.

Parameter	Abbreviations	Standard Method
pH	pH	SM 4500-H+
Nitrate	NO_3^-	SM 4500- NO3- E
Phosphorus	P	SM 4500-P B & E
Temperature	T	SM 2550 B
Salinity	Sal	SM – 2520-B
Total solids	TS	SM 2540D
Dissolved oxygen	DO	SM 4500-O G
Biochemical oxygen demand	BOD_5	SM 4500-O G
Total coliforms	TC	SM 9222B
Ammonium	NH_4^+	SM 4500-NH3-C

2.4. Data treatment

The similarities between sampling stations were studied using Kruskal-Wallis test, due to the non-normality of the collected data and the presence of outliers in the data base. Kruskal-Wallis is a non-parametric measure of the correlation between variables, that uses the median and not the mean to compare two or more samples(Pati, Dash, Mukherjee, Dash, & Pokhrel, 2014)(Wilbers, Becker, Nga, Sebesvari, & Renaud, 2014). Principal component analysis was used to find new variables represented by a linear combination of variables having correlations via the variance-covariance matrix of several multivariate variables; it explains most of the total variations with some important principal components(Fan, Cui, Zhao, Zhang, & Zhang, 2010)(Al-Mutairi, Abahussain, & El-Battay, 2014). All mathematical and statistical calculations were implemented using Statgraphics 16.1.11 and Microsoft Office Excel 2015.

3. RESULTS AND DISCUSSION

The data obtained for this research during the period 2001-2014 was compared with the established targets for BOD5, OD, phosphorus, total coliforms, etc. in Colombian regulations. allowed values are in table 2.

Table 2. Allowed values for water quality parameters.

Parameter	Allowed value
$P(mgL^{-1})$	< 0.003
T(°C)	27 – 30
Sal	33 – 36
$TS(mgL^{-1})$	≤ 90
$DO(mgL^{-1})$	≥ 4
$BOD_5(mgL^{-1})$	≤ 3
TC(NMP)	≤ 5000
NO_3^- (mgL^{-1})	≤ 5
pH	6.5 – 8.5
$NH_4^+(mgL^{-1})$	≤ 1

The behavior of BDO_5 for the sampling stations is depicted in Fig 2., it can be seen a great number of outliers which demonstrate the non-normality of the data. Kruskal-Wallis test showed no statistical difference in the stations for this parameter (Statistical = 1,89 p-value = 0,39).

Lequerica, Matunilla and Baru showed outliers beyond the values allowed by the law, in this way none of the stations set with the law for BDO_5.(Ministry of Environment Housing and Territorial Development, 2015).

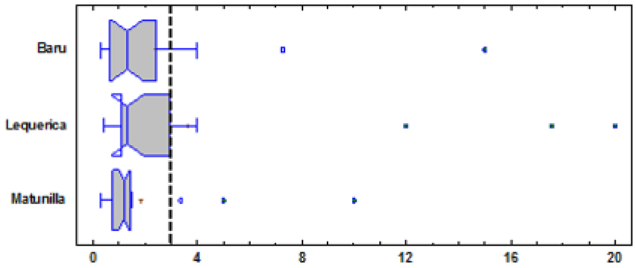

Fig. 2. BDO5 Box-plot.

For total coliforms fig 3. Shows that there is a significant difference between Lequerica and Matunilla with Baru values for this variable (Statistical = 29,51 p-value = 3,89E-7), these differences are notorious in rainy season due to the rise of the discharges of microorganism in the two canals of the canal del Dique.

According to table 2. Lequerica and Matunilla do not set with law levels for this variable. the high values found in lequerica and matunilla represent a threat to human and animal health.

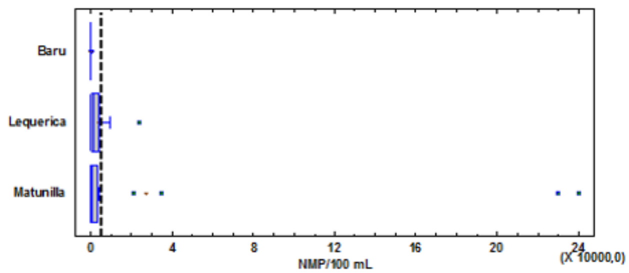

Fig. 3. Total coliform Box-plot.

Total solids in Lequerica y Matunilla has no significant difference between then, but Baru had a different behavior with the others sampling points. It showed outliers that may be a threat for the economic activity in this zone related to tourism. Also, the coral reef may be influenced by this parameter, because excess of solids may lead to disease in the corals. Lequerica and Matunilla do not set with the law for total solids. This Value must be less than 90 mg/L (black line in fig.4) according to Colombian law and the values are near to 1500 mg/L in Lequerica and Matunilla.

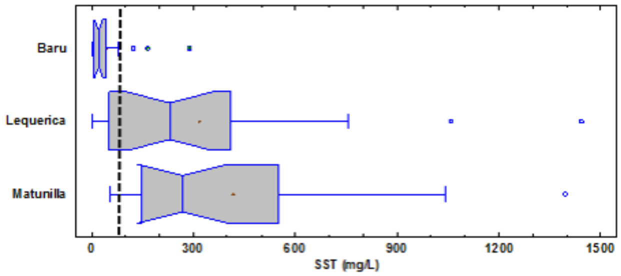

Fig. 4. Total solids Box-plot.

Fig. 5 shows the variation of salinity in the sampling points, as showed in figure Baru station had bigger values for this measured variable, it also had a significant difference with the two others points. Kruskal wallis test presented a Statistical equal to 18,24 and a p-value of 0,0001 which probed the difference between the stations. This is result of the mix between marine water and fresh water coming from the Magdalena river. Low salinity represents also a threat for coral reef health.(Fabricius, 2005).

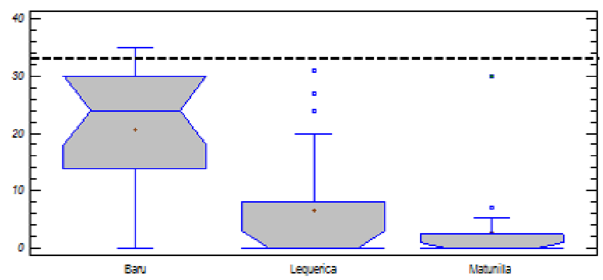

Fig. 5. Salinity Box-plot.

Nitrates range was [0,012-0,833] mg/L for Baru station, Lequerica was [0,012-3,384] and Matunilla was [0,012-15,1], which means that only Matunilla had points out of the allowed range for this variable, as shown in fig. 6. Nitrate is the final product of nitrification process; it is harmful for humans and harmless for fishes and others aquatic organisms.(Cárdenas Calvachi & Sánchez Ortiz, 2013).

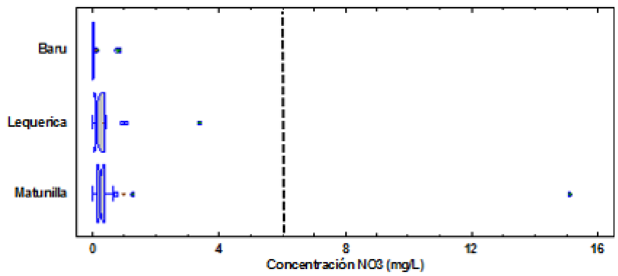

Fig. 6. Nitrates Box-plot.

According to fig. 7 there is not a statistical difference among the three sampling points for phosphorus behavior with a 95% of confidence. It may be seen that Lequerica and Matunilla have more variance in this values than Baru. This is probably cause for excess of fertilizer used by the agricultural activities and waste water effluents. None of the points set with regulatory values for phosphorus(Fabricius, 2005).

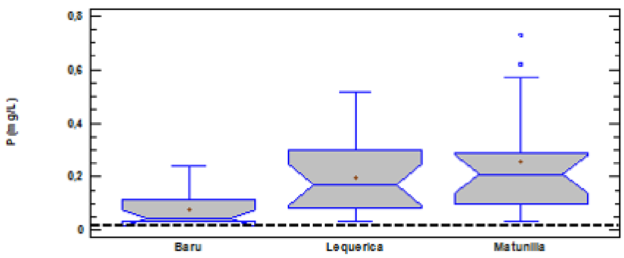

Fig. 7. Total Phosphorus Box-plot.

All three stations showed no statistical differences in ammonium concentrations, Matunilla had a wider range, that could be result of rainy season or waste water from the different anthropogenic activities around the bay, the three stations had good results for this parameter.

It has been demonstrated that nutrient enrichment (ammonium, phosphorus and nitrates) has a negative effect on coral reproduction and its calcification, decreasing the growth rate of the reef. This may bring to algal overgrowth of the reef, with a phase shift coral/macroalgae. Nutrient enrichment is also correlated with a major incidence in coral diseases(Gavio, Palmer-Cantillo, & Mancera, 2010).

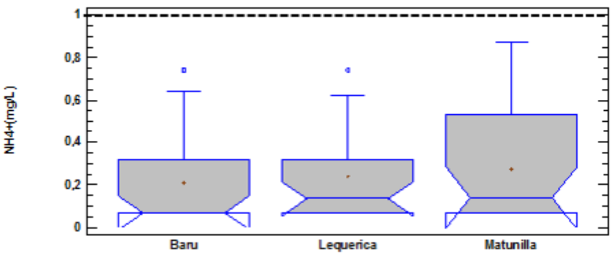

Fig. 8. Ammonium Box-plot.

3.1. Principal component analysis

With a large number of variables under study, it is difficult to study and interpret the dispersion matrix in an appropriate way. To interpret the data in a more meaningful way it is necessary to reduce the number of variables to a smaller number, without the need to omit information; This objective is achieved by applying the principal component analysis method.(Taoufik, Khouni, & Ghrabi, 2017).

To identify the number of principal components, the scree plot was used for the data in dry and rainy season. Fig. 9. shows an evident change in the slope in the fourth component, so that three components are conserved, which have eigenvalues greater than unity and explain the 68.81% of the variance of the information contained in the initial database.

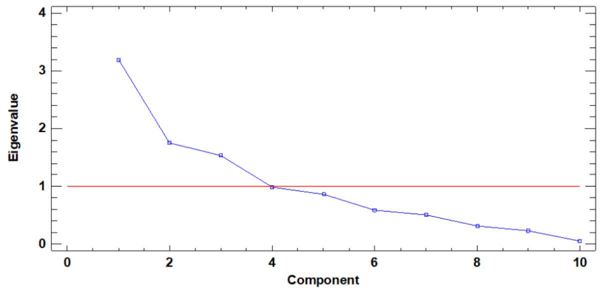

Fig. 9. Scree plot of eigenvalue in rainy season.

In this study the number of data is more than 250, according to this the accepted loadings are those higher than 0,40 (Hair, Black, Babin, & Anderson, 2010). Loading of the three conserved components are presented in table 3. PC1 explains 31,92% of the variance and is strong positive influenced by variables TS and pH with a negative contribution, this factors mainly represent the influx of sediments to the bay and pH may be caused by the inflow of industrial wastewater. PC2 explains 17,56% of the variance associated with NO3-, T and TC, this factor is an anthropogenic factor and may be related to wastewater and over fertilization. PC3 explain 15,3% and variables NH4+, NO3-, DO have significant loadings, this factor is possible linked to fertilizers used in shrimp factories near the bay, the negative loading for OD relates the inverse relation between OD and nitrates.(Chow et al., 2016)

Table 3. Loadings of variables on three principal components during rainy season.

	Component 1	Component 2	Component 3
NH_4^+	0,267	0,009	**-0,589**
P	0,393	-0,223	0,249
NO_3^-	0,130	**0,435**	**0,399**
DO	-0,283	0,030	**-0,446**
pH	**-0,505**	0,062	-0,083
T	-0,200	**0,500**	0,256
TS	**0,426**	0,197	-0,298
Sal	-0,282	-0,210	0,013
TC	0,199	**0,567**	-0,172
BDO_5	-0,278	0,314	-0,200
Var%	31,927	17,577	15,303
Cum%	31,927	49,504	64,807

For Dry season the results of the number of eigenvalues are shown in Fig. 10. This figure shows that 4 components have eigenvalues greater than unity and explain 69.6% of the information variance.

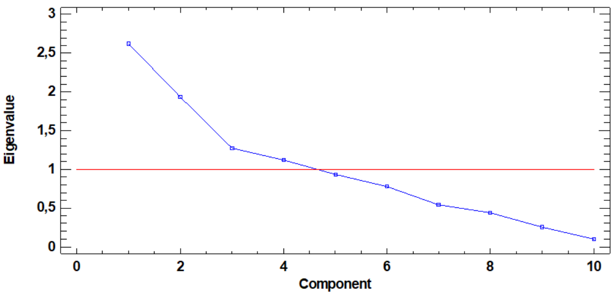

Fig. 10. Scree plot of eigenvalue in dry season.

Loading of the four conserved components are presented in table 4. PC_1 explains 26,28% of the variance and is strongly influenced by variables TS, DO, Sal with a negative contribution. PC_2 explains 19,38% of the variance associated with NO3- BDO5, pH in a negative way. PC_3 explain 12,73% and variables TC and T with a negativity significant loading, this factor is the input from anthropogenic (rural sewage and agricultural) and natural sources. PC_4 has a substantial negative contribution from NH_4+ and a minor contribution from TC, it explains 11,23 % of the variance in data and could be an anthropogenic contribution coming from sewage.

Table 4. Loadings of variables on four principal components during dry season.

	Component1	Component2	Component3	Component 4
NH_4^+	0,167	-0,250	-0,123	**-0,631**
P	0,381	-0,339	0,260	-0,198
NO_3^-	0,115	**0,457**	0,215	-0,221
DO	**-0,459**	-0,085	0,129	-0,060
pH	-0,209	**-0,523**	-0,186	-0,080
T	0,294	0,146	**-0,566**	0,378
TS	**0,512**	0,099	-0,168	-0,021
Sal	**-0,437**	0,152	-0,287	0,042
TC	0,092	-0,074	**0,610**	**0,463**
BOD_5	-0,098	**0,520**	0,118	-0,377
Var%	26,276	19,367	12,729	11,232
Cum%	26,276	45,644	58,372	69,604

CONCLUSIONS

Due to the high number of outliers in data for DOB5, none of the points set with the recommended values, for example Lequerica has outlier near to 20 mg/L. Total coliform range is [1.8-240,000] for Matunilla and [150-24,000] for Lequerica, these values are a real threat for any living organism in this area. Salinity levels are low [0-30] in the bay due to the influx of fresh water from the river, this is a vital factor to corals because the need to have a high salinity to stay healthy. [0,031-0,73] and [0,08-0,52] are the ranges of phosphorus for Matunilla and Lequerica respectively, the high values of phosphorus encourage the growth of macroalgaes over coral. Values of phosphorus and salinity behavior suggest problems in future health of coral reefs near this bay, since high values of phosphorus and low values of salinity contribute to the deterioration of this species. Also, total coliforms outliers may become a dangerous factor for flora, fauna and inhabitants of the region.

In this study, different multivariate statistical techniques were used to evaluate variations in surface water quality of Barbacoas Bay. Kruskal wallis Analysis showed similar water quality characteristics for Matunilla and Lequerica for almost all the time period. Based on the information obtained from this study, it is possible to exclude total phosphorus of the monitoring procedure because of its loadings in principal components analysis (<0.4) and this will lead to lower associated costs.

PCA helped to identify the sources responsible for water quality variables. The main cause of degradation to the bay is determined to be the discharge agricultural wastes, domestic sewage water from Matunilla and Lequerica canals. This study illustrates the usefulness of multivariate statistical techniques in the analysis and interpretation of complex data sets, in identifying pollutant sources, and in understanding variations in water quality for effective water management. These results should be considered for future planning and management of the bay.

REFERENCES

Acosta, J. C., & Baldiris, I. (2015). Analysis of water quality variation in barbacoas bay – Cartagena during the period 2001-2014. *Revista Ingeniería E Innovación*, 3, 7–17.

Al-Mutairi, N., Abahussain, A., & El-Battay, A. (2014). *Spatial and temporal characterizations of water quality in Kuwait Bay. Marine Pollution Bulletin*, 83(1), 127–131. https://doi.org/10.1016/j.marpolbul.2014.04.009

Alberto, W. D., María del Pilar, D., María Valeria, A., Fabiana, P. S., Cecilia, H. A., & María de los Ángeles, B. (2001). Pattern Recognition Techniques for the Evaluation of Spatial and Temporal Variations in Water Quality. A Case Study: Suquía River Basin (Córdoba–Argentina). *Water Research*, 35(12), 2881–2894. https://doi.org/10.1016/S0043-1354(00)00592-3

American Public Health Association, American Water Works Association, & Water Environment Federation. (1999). *Standard Methods for the Examination of Water and Wastewater* (20th ed.). Washington, DC.

Bhuiyan, M. A. H., Rakib, M. A., Dampare, S. B., Ganyaglo, S., & Suzuki, S. (2011). Surface water quality assessment in the central part of Bangladesh using multivariate analysis. KSCE *Journal of Civil Engineering*, 15(6), 995–1003. https://doi.org/10.1007/s12205-011-1079

Cárdenas Calvachi, G. L., & Sánchez Ortiz, I. A. (2013). Nitrógeno en aguas residuales: orígenes, efectos y mecanismos de remoción para preservar el ambiente y la salud pública. *Universidad Y Salud*, 15(1), 72–88. Retrieved from http://revistas.udenar.edu.co/index.php/usalud/article/view/375

Chow, M. F., Shiah, F. K., Lai, C. C., Kuo, H. Y., Wang, K. W., Lin, C. H., … Ko, C. Y. (2016). Evaluation of surface water quality using multivariate statistical techniques: a case study of Fei-Tsui Reservoir basin, Taiwan. *Environmental Earth Sciences*, 75(1), 1–15. https://doi.org/10.1007/s12665-015-4922-5

Fabricius, K. E. (2005). Effects of terrestrial runoff on the ecology of corals and coral reefs: Review and synthesis. *Marine Pollution Bulletin*. https://doi.org/10.1016/j.marpolbul.2004.11.028

Fan, X., Cui, B., Zhao, H., Zhang, Z., & Zhang, H. (2010). Assessment of river water quality in Pearl River Delta using multivariate statistical techniques. Procedia Environmental Sciences, 2, 1220–1234. https://doi.org/10.1016/j.proenv.2010.10.133

Gavio, B., Palmer-Cantillo, S., & Mancera, J. E. (2010). Historical analysis (2000-2005) of the coastal water quality in San Andr??s Island, SeaFlower Biosphere Reserve, Caribbean Colombia. *Marine Pollution Bulletin*, 60(7), 1018–1030. https://doi.org/10.1016/j.marpolbul.2010.01.025

Giraldo, A. G., Osorio, A. F., Toro, F. M., Osorio, J. D., Oscar, a, & Arrieta, A. (2009). Patrón de circulación en Bahía Barbacoas y su influencia sobre el transporte de sedimentos hacia las islas del Rosario. Avances En …, 21–40. Retrieved from http://www.bdigital.unal.edu.co/4766/

Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis* (7th ed.). Prentice hall.

Ministry of Environment Housing and Territorial Development. (2015). *Resolution 0631. Resolution 0631 of march 17- 2015*. Bogota. Retrieved from http://www.fenavi.org/images/stories/estadisticas/article/3167/Resolucion_0631_17_marzo_2015.pdf

Moreno-Madriñán, M. J., Rickman, D. L., Ogashawara, I., Irwin, D. E., Ye, J., & Al-Hamdan, M. Z. (2015). Using remote sensing to monitor the influence of river discharge on watershed outlets and adjacent coral Reefs: Magdalena River and Rosario Islands, Colombia. *International Journal of Applied Earth Observation and Geoinformation*, 38, 204–215. https://doi.org/10.1016/j.jag.2015.01.008

Pati, S., Dash, M. K., Mukherjee, C. K., Dash, B., & Pokhrel, S. (2014). Assessment of water quality using multivariate statistical techniques in the coastal region of Visakhapatnam, India. *Environmental Monitoring and Assessment*, 186(10), 6385–6402. https://doi.org/10.1007/s10661-014-3862-y

Pinilla, G., Gutiérrez, Á., & Ulloa, G. (2007). *Efectos ecológicos de la derivación de aguas y sedimentos hacia la bahía de barbacoas*, (1033), 1–42.

Restrepo, J. D., Park, E., Aquino, S., & Latrubesse, E. M. (2016). Coral reefs chronically exposed to river sediment plumes in the southwestern Caribbean: Rosario Islands, Colombia. *Science of the Total Environment*, 553, 316–329. https://doi.org/10.1016/j.scitotenv.2016.02.140

Taoufik, G., Khouni, I., & Ghrabi, A. (2017). Assessment of physico-chemical and microbiological surface water quality using multivariate statistical techniques: a case study of the Wadi El-Bey River, Tunisia. Arabian *Journal of Geosciences*, 10(7), 181. https://doi.org/10.1007/s12517-017-2898-z

Wilbers, G.-J., Becker, M., Nga, L. T., Sebesvari, Z., & Renaud, F. G. (2014). Spatial and temporal variability of surface water pollution in the Mekong Delta, Vietnam. *Science of The Total Environment,* 485–486, 653–665. https://doi.org/10.1016/j.scitotenv.2014.03.049