Modeling of Suches River Contamination Using Water Quality Indices, Puno - Peru, 2009-2014

Teófilo Donaires Flores
Facultad de Ingeniería Química
Universidad Nacional del Altiplano Puno-Perú
tdonaires@unap.edu.pe

Accepted 6th October 2017

Abstract

The Suches River binational axis of socioeconomic development in the study area, informal mining as a whole produces a series of gaseous and solid pollutants, liquids, which alter the atmosphere and environment. The statistical method used to validate models pollution is the Multivariate Analysis, which aims to simultaneously analyze multivariate datasets formed by OD, pH, BOD 5, nitrates, phosphates, temperature, turbidity, total solids, arsenic, cadmium, copper, iron, mercury and lead. It was established during the 2009-2014 and the three weather conditions. The Suches River in the first part presents quality index 50.00 according ICASUCHES is considered acceptable quality, the model is $Y = -134.8 - 0.1138X - 0.270X_2 - 4.25X_3 - 1.001X_10 - 0.207X_{11} + 0.622X_2 - 1.107X_{13} \pm \varepsilon$. $R^2$ 73.50%, presents a pollution index 0.20 represents low contamination presence of micro pollutants; the intermediate portion 36.00 presents quality index representing poor quality, where it reaches the further deterioration of the quality of the river, the model is $Y = -134.8 - 0.1138X - 0.0884X_2 + 1.855X_4 + 0.7470X_5 - 0.4605X_10 + 0.1440X_{11} - 0.5409X_{13} + 0.2422X_{14} \pm \varepsilon$. $R^2$ 98.07%, 0.70 presents pollution index representing high contamination presence of heavy metals; Is finally mouth of Lake Titicaca has 42.00 index represents a slight recovery in quality, although in this section indices classify water between inadequate and acceptable quality, the model is $Y = 123.2 + 0.241X_5 + 0.527X_{13} \pm \varepsilon$. $R^2$ which has 50.00%, 0.50 presents pollution index representing median presence of mercury contamination.

Keywords: Quality, Pollution Index, Multivariate, Suches River.

Introduction

Water is an indispensable and fundamental resource for living things, especially for humans, making it the most abundant liquid and the most important resource on earth. Water, present in different forms and used for various purposes, occupies a high proportion in relation to the earth’s surface. It is present in the seas and oceans, in surface waters and in groundwater, and can be used, among other things, for domestic, industrial and agricultural tasks. The survival of man as a species is due in part to the use of natural resources; however, the lack of planning and the lack of knowledge of the possible consequences of a poor use brought about the contamination of the natural environment. The water resource (necessary for life) has been strongly affected by substances that are increasingly aggressive and difficult to treat because of their chemical nature of substances present in waste falling into streams. Physico-chemical parameters provide extensive information on the nature of water chemical species and their physical properties, without providing information on their influence on aquatic life; (Orozco, Perez, Gonzalez, & Alfayate, 2002). In addition, the use of both methods in the evaluation of the water resource is important. The treatment of the data obtained in monitoring is often an expensive and often difficult task for the different actors involved in the process of quality assessment, because at present the values obtained must be able to solve different types of conflicts such as the use of water and the ecological integrity of aquatic systems, which also involve socioeconomic aspects (Fernandez & Solano, 2005).

The implementation of new methodologies that involve more than two parameters for the assessment of water quality is becoming increasingly important, water quality indices encompass several parameters, mostly physicochemical and in some cases microbiological, that allow the reduction of information to a simple expression, known as: water quality indexes (ICA) and water pollution indexes (ICO). The results of a monitoring should allow to solve different types of conflicts like the use of the water and the ecological integrity of the aquatic systems, which involve socioeconomic aspects, reason why the ICA and ICO are an important tool since its calculation involves more than one variable, so that the correct use of these indicators allows them to be used for the evaluation of water resources management programs (Fernandez & Solano, 2005).

The Suches river basin is of Peru - Bolivia binational character, located in the district of Cojata, Huancané province, department and region of Puno, where different environmental pollution processes are generated, deteriorating the environmental quality of the Lake Titicaca basin, because it ends, affecting the living conditions of the communities involved both abiotic and biotic. According to studies carried out by a group of researchers from the National University of the Altiplano, they have found heavy metals of 0.0010 mg / L of arsenic, 0.0200 mg / L of cadmium, 0.0840 mg / L of copper, 0.4933 mg / L of iron, 0.0002 mg / L of mercury and 0.0500 mg / L of lead (Valdivia, et al., 2012).

At present, several methodologies have already been developed in other countries, with the purpose of preparing indexes that can characterize, in a simple value, the water quality according to the most critical constituents of knowledge of the possible consequences of a poor use...
recognized. As these studies have evolved, new techniques such as those mentioned by Dinius (1987) have been identified, which make this tool a more complete and safe tool for the management of potable water supplies. All these techniques and methodologies will be evaluated and taken into account when designing the Water Quality Indexes, in order to achieve a more representative value to the existing reality in the rivers. The concrete determination of the quality index models of a natural system will depend on the type of pollutant and the type of natural system under study. The results of the modeling will allow to know, for each pollutant: spatial and temporal distribution of the concentration in the natural environment as a consequence of a concrete series of discharges fraction of the discharge that disappears from the processes of transformation and transportation. (Piller, Tavard, & Tavaret, 2014).

The general objective is to provide the authorities responsible for the management and administration of the water resource with a modern, efficient and reliable tool that allows the entity to analyze and make decisions regarding activities, works, impacts, current status and future of the River. This tool deals with technical aspects related to hydrology of the river basin, hydraulics and sedimentology, the geomorphological, morphological, morphodynamic and water quality characteristics of the Suches river. Water quality indexes are an effective tool for the management and management of surface water flows in watersheds (Marquez C., 2008).

**Materials and Methods**

**Information on Historical Water Quality Monitoring Data**

It contains the results of the evaluation of the physical, chemical and microbiological parameters related to the water quality in the area of Suches rivers in the Peruvian side, corresponding to the transition period of avenue in 2009 and 2010 (Valdivia, and others, 2012), information is also available on the monitoring carried out by the Nacional del Agua (ANA) also known as National Water Agency based in Puno (ANA, Technical Report 023 - 2013 Results monitoring water quality in the Suches River Basin, 2013 ), also has geo referenced the monitoring points, are shown in table 3. These coordinates are the sampling points made by the ANA. Historical data were used to model pollution of the Suches river.

The statistical method used for the validation of pollution models of the Suches river is the Multivariate Analysis. One of the most useful statistical methods, whose purpose is to simultaneously analyze multivariate data sets in the sense that there are several variables measured for each individual or object studied in the pollution of the Suches river. Its purpose lies in a better understanding of the phenomenon under study by obtaining information that univariate and bivariate statistical methods are unable to achieve (Hair, Anderson, Tatham, & Black, 1999).

**Multiple Linear Regression**

In the multiple linear regression we try to determine the relationship between the dependent variable (Y) and two or more independent variables (X1, X2, X3, ..., XK) also called regressor variables. In this case the dependent variable is affected by the changes that are made to the independent variables as a whole.

The relationship between the regressor variables and the dependent variable is established by the general multiple linear regression model (Montgomery, 2004).

\[
Y = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \beta_3 X_3 + \cdots + \beta_k X_k
\]

where \(\beta_0, \beta_1, \beta_2, \ldots, \beta_k\) are the parameters of the model (we have k independent variables and p parameters).

### Table 1. Operationalization of variables

<table>
<thead>
<tr>
<th>Variables</th>
<th>Description</th>
<th>Indicators</th>
<th>Units or Categories</th>
<th>Valuation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dissolved oxygen (X1)</td>
<td>Concentration</td>
<td>mg/L</td>
<td>&lt; 5 mg/L</td>
<td>affect the biological diversity</td>
</tr>
<tr>
<td>pH (X2)</td>
<td>Concentration</td>
<td>Units</td>
<td>6.0 8.5 typical value</td>
<td>in surface water</td>
</tr>
<tr>
<td>BOD5 (X3)</td>
<td>Concentration</td>
<td>mg/L</td>
<td>&gt; 10 mg/l</td>
<td>waters impacted by discharges of waste water, particularly near the point of discharge</td>
</tr>
<tr>
<td>Nitrate (X4)</td>
<td>Concentration</td>
<td>mg/L</td>
<td>&lt; 0.01 ppm</td>
<td></td>
</tr>
<tr>
<td>Phosphate (X5)</td>
<td>Concentration</td>
<td>mg/L</td>
<td>&lt; 0.01 ppm</td>
<td></td>
</tr>
<tr>
<td>Temperature (X6)</td>
<td>Units</td>
<td>°C</td>
<td>units</td>
<td></td>
</tr>
<tr>
<td>Turbidity (X7)</td>
<td>Concentration</td>
<td>NTU</td>
<td>&lt; 1000 typical value</td>
<td>in surface water</td>
</tr>
<tr>
<td>Total solids (X8)</td>
<td>Concentration</td>
<td>mg/L</td>
<td>110 world average in surface water</td>
<td></td>
</tr>
<tr>
<td>Arsenic (X9)</td>
<td>Concentration</td>
<td>mg/L</td>
<td>&lt; 0.01 ppm</td>
<td></td>
</tr>
</tbody>
</table>
Cadmium ($X_{10}$) Concentration mg/L < 0.01 ppm
Copper ($X_{11}$) Concentration mg/L < 0.01 ppm
Iron ($X_{12}$) Concentration mg/L < 0.01 ppm
Mercury ($X_{13}$) Concentration mg/L < 0.01 ppm
Lead ($X_{14}$) Concentration mg/L < 0.01 ppm

### Modeling ICA-SUCHES

<table>
<thead>
<tr>
<th>Water Quality (Y)</th>
<th>Water Quality Index</th>
</tr>
</thead>
<tbody>
<tr>
<td>Index Scale</td>
<td>Classification</td>
</tr>
<tr>
<td>90-100</td>
<td>Optimal quality</td>
</tr>
<tr>
<td>70-90</td>
<td>Good quality</td>
</tr>
<tr>
<td>50-70</td>
<td>Acceptable quality</td>
</tr>
<tr>
<td>25-50</td>
<td>Poor quality</td>
</tr>
<tr>
<td>0-25</td>
<td>Very bad quality</td>
</tr>
<tr>
<td>Source: (NSF N.S., 2003)</td>
<td></td>
</tr>
</tbody>
</table>

### Calculation of ICAs and ICOs

Taking into account that the methodology used in the estimation of the multiplicative ICAs to be applied in this study is based on the geometric calculation of the values found by the subscripts ($I_i$) and the assigned weight ($W_i$) for each parameter used in each index in particular, previously selected ICAs were calculated.

\[
ICA_{SUCHES} = \sum I_i W_i
\]

\[
ICA = \prod_{i=1}^{n} I_i^{W_i} = (I_1^{W_1})(I_2^{W_2})(I_3^{W_3})...I_n^{W_n}
\]

In the case of the ICOs, a calculation procedure based on the arithmetic mean of the values found in each of the sub-indexes ($I_i$) estimated for each physicochemical and microbiological parameter evaluated was used. In order to determine the value of ICAs and ICOs in the Suches river, an electronic spreadsheet was developed that could be applied to any date and monitoring station and that allowed to reduce the calculation time of the indices. In this electronic sheet were introduced the equations that determine the value of the subscript of each parameter, the respective weights and the corresponding calculation equation of each index.

### Results and Discussions

#### Section I. Exit of Suches Lake - Hito 19

The values found of the application of the ICES of CETESB in Section I the river presents its best characteristics for its destination for human consumption with ICA values that classify its waters as of a regular quality to poor quality.

Regarding the ICA values of Rojas presented in Section I, it can be observed that during the years 2009-2014 and the three climatic conditions studied, most of the stations located in this section report acceptable and inadequate water quality with a tendency towards progressive deterioration.

This demonstrates the importance of mitigating and formalizing informal mining. In general, according to the behavior of the four ICAs applied to the Suches river, it is found that in Section I the water quality is presented in the condition of transition to inadequate quality in which the flow rate allows damping the deterioration of the quality by dilution, also the contents of material in suspension in the current do not reach to increase enough to reduce significantly the value of the ICAs.

In summer conditions (avenue epoch), in spite of a better dilution due to higher flow rates, the concentrations of suspended solids are also increased, thus decreasing the water quality, which contributes to the contamination of the Suches river.

Section I presents an ICOMI value of 0.28, water is classified with a quality that makes agricultural use unprotected without treatment due to the presence of heavy metals. Therefore, it requires preliminary treatment for agricultural use. The ICOMO of 0.20, classifies the water with a quality that makes risky the agricultural use without treatment by the presence of heavy metals.

Therefore, it requires preliminary treatment for agricultural use. ICOSUS values are presented with a maximum value of 0.64 and a minimum value of 0.42, which is high contamination.

Interpretation: The step-by-step method ends in 6 steps. The first variable eliminated from the model is the total solids index ($X_8$), whose test value $t$, 0.13, is the smallest of all, then eliminated, lead index ($X_{14}$), turbidity index ($X_7$), arsenic index ($X_9$), phosphate index ($X_5$), dissolved oxygen index ($X_1$) and temperature index ($X_6$) in that order.

The Model with Predictor Discrimination:

\[
Y = 579 + 0.245 X_2 - 0.270 X_3 - 4.25 X_4 - 1.001 X_{10} - 0.207 X_{11} + 0.622 X_{12} - 1.107 X_{13} ± e
\]

Which has a $R^2$ of 73.50%, while the complete model with 14 predictor variables has a $R^2$ of 91.30%, has lost a 17.80% confidence in the predictions, this indicates us with greater number of predictor variables can be modeled with greater approximation to reality.

A good approximation can be seen in the residual diagram.
**Interpretation:** there is normality in the distribution of errors, because the points do not move far from a straight line on the first plot and some symmetry that can be seen in the second. However, the presence of the outliers at both ends is clear, which affects the normality condition.

The plot of residuals versus the order of observation and plot of residual versus predicted values ("fits"). suggests that observations 3, 4, 13, 16 and 17 are outliers in the vertical sense, these outliers. The plot of residuals versus predicted values suggests that the variance of errors is constant because there is no definite pattern they follow the points, therefore, it occurs randomly.

**Section II. Hito 19- Puente Castilla**

In this section there is a significant decrease between the beginning (Milestone 19) and the end of it (Puente Castilla) (Approximately 15% in the ICA NSF and Dinius and 10 to 25% in the CETESB and Rojas indexes).

This reduction can be associated with the influence on the water quality of the discharges from the informal mines of the Lunar de Oro, Trapiche and dispersed areas of the urban area, the mining wastewater, as well as the discharges of some the tributary rivers to the river Suches, which receive mining, agricultural and domestic wastewater from different communities located on the banks of the Suches river. Likewise, the bofedales of zone also present great deterioration, which directly affects the cattle of the zone of study.

Section II presents ICOMI values, reaching maximum values of 0.68 and minimum values of 0.50, which indicates that the water is contaminated by the mining activity that is developed, particularly by the presence of heavy metals.

ICOSUS, a high contamination is observed because it has maximum values of 0.80 which indicates high contamination due to the suspended solids.

**Interpretation:** The method ends in 6 steps. The first variable eliminated from the model is the BOD5 (X3), whose test value t, 0.02, is the smallest of all, then removed, turbidity index (X7), temperature index (X6), index of arsenic (X9), total solids index (X8) and iron index (X12) in that order.

**The Predictor Discrimination Model:**

\[ Y = -134.8 - 0.11138 X_1 - 0.0884 X_2 + 1.855 X_4 + 0.4750 X_5 - 0.4605 X_{10} + 0.1440 X_{11} - 0.5409 X_{13} + 0.2422 X_{14} + \varepsilon \]

Which has a R2 of 96.92%, while the complete model with 14 predictor variables has a R2 of 98.07%, has lost 1.15% of predictability, this indicates us with more number of predictor variables can be modeled with greater approximation to reality.

**Interpretation:** there is normality in the distribution of errors, because the points do not move far from a straight line on the first plot and some symmetry that can be seen in the second.

However, the presence of the outliers at both ends is clear, which affects the normality condition. The plot of residuals versus the order of observation and plot of residual versus predicted values ("fits"). suggests that observations 2, 10 and 13 are outliers in the vertical sense, these outliers.

The plot of residuals versus predicted values suggests that the variance of errors is constant because there is no definite pattern following the points.
Section III. Puente Castilla - Mouth Lake Titicaca

Section III presents approximate values to those determined in Section II, showing a slight increasing tendency especially in the last stations located in that section. However, despite this recovery, the water of the Suches river is still within the lowest classification strips presented by the different ICAs applied (Poor quality according to ICA - NSF and CETESB ICA in most stations, very poor to inadequate quality according to Rojas and waters that necessarily require treatment according to the ICA of Dinius.), therefore, it is not suitable for human consumption, nor consumption of animals, nor for irrigation.

Section III presents an ICOMI of maximum of 0.42 and a minimum value of 0.30 at the mouth of the Suches river, which indicates that during the period 2009-2014 the climatic conditions of summer are contaminated, where water quality is classified as medium to high pollution, presents high irrigation for agricultural use. The ICOMO maximum of 0.18 and a minimum value of 0.10 at the mouth of the river Suches, which indicates low contaminated organic matter, during the period 2009-2014, the climatic conditions of summer, where the quality of the water is classified as low pollution by the presence of organic matter. ICOSUS presents maximum values of 0.70 and minimum values of 0.60 which classifies high contamination.

Interpretation: The method ends in 6 steps. The first variable removed from the model is the turbidity index (X8), whose test value t, 0.13, is the smallest of all, then removed, lead index, turbidity index, arsenic index, index of cadmium, copper index and iron index, in that order. The Model with Predictor Discrimination: \[ Y = 12.3 + 0.241X5 + 0.527X13 \pm \varepsilon \]

Technically this model has no non-significant, which has a R2 of 50.08%, while the complete model with 14 predictor variables has a R2 of 90.10%, has lost a 40.02% confidence in the predictions, this indicates us with greater number of predictor variables can be modeled with greater approximation to reality.
Interpretation: there is normality in the distribution of errors, because the points do not move far from a straight line on the first plot and some symmetry that can be seen in the second.

However, the presence of the outliers at both ends is clear, which affects the normality condition. The plot of residuals versus the order of observation and plot of residual versus predicted values ("fits"). suggests that observations 4 and 10 are outliers in the vertical sense, these outliers.

The plot of residuals versus predicted values suggests that the variance of errors is constant because there is no definite pattern following the points.

<table>
<thead>
<tr>
<th>Years</th>
<th>Station</th>
<th>Physicochemical Index</th>
<th>ICASUCHES</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009-2014</td>
<td>Section I</td>
<td>50·00</td>
<td>Acceptable Quality</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Section II</td>
<td>36·00</td>
<td>Poor quality (very bad)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Section III</td>
<td>42·00</td>
<td>Poor quality (contaminated)</td>
<td></td>
</tr>
</tbody>
</table>

Source: self made

As it can be shown the greater contamination is in section II, this is due to the mining activity that is developed in Lunar de Oro of the Suches river basin.

<table>
<thead>
<tr>
<th>Years</th>
<th>Station</th>
<th>Physicochemical Index</th>
<th>ICASUCHES</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>2009-2014</td>
<td>Section I</td>
<td>0·20</td>
<td>Low pollution</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Section II</td>
<td>0·70</td>
<td>High pollution</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Section III</td>
<td>0·50</td>
<td>Medium contamination</td>
<td></td>
</tr>
</tbody>
</table>

Source: self made

As it can be shown the greater contamination is in section II, this is due to the mining activity that is developed in Lunar de Oro of the Suches river basin.

Conclusions

A bibliographic review has been carried out for the modeling of the Water Quality Indexes (ICA). The National Water Sanitation Index (ICA-NSF), the Water Quality Index of CETESB (ICA-CETESB) (ICA-Rojas), the Dinius Water Quality Index (ICA-Dinius) and the Pollutant Indices (ICOs), formed by the Mineralization Contamination Indices (ICOMI), Suspended Solids Contamination Index (ICOSUS) and Organic Matter Contamination Index (ICOMO). The evaluation of the water quality of the Suches river by means of these Water Quality Indexes (ICA) and pollution (ICO), systematically show three sections (exit from Suches Lake - Hito 19, Puente Castilla - mouth of Lake Titicaca). The models obtained in the two conditions are: Section I The Model with Predictor Discrimination

\[
Y = 703 - 0.110 X_1 + 0.331 X_2 - 0.426 X_3 - 5.24 X_4 + 0.218 X_5 + 0.145 X_6 - 1.099 X_10 + 0.2264 X_11 + 0.562 X_12 - 1.558 X_13 \pm \epsilon
\]

Which has a R² of 86.08%, while the complete model with 14 predictor variables has a R² of 91.30%, a loss of 5.22% in predictions has been lost. predictor variables can be modeled with greater approximation to reality. In addition, it has a pollution index of 0.20 representing low pollution. Quality rating of 50 representing acceptable quality. Section II.

The Model with Predictor Discrimination

\[
Y = -174.7 - 0.1116 X_1 - 0.1503 X_2 + 2.285 X_4 + 0.5723 X_5 - 0.0877 X_8 - 0.4288 X_10 + 0.1525 X_11 - 0.0825 X_12 - 0.5100 X_13 + 0.2421 X_14 \pm \epsilon
\]

Which has a R² of 97.84%, while the complete model with 14 predictor variables has a R² of 98.07%, has lost a 0.23% confidence in the predictions, this indicates us with greater number of predictor variables can be modeled with greater approximation to reality. It presents a contamination index of 0.70 which represents high pollution, and effectively is the zone of greater mining activity. Quality index of 36.00 representing poor quality as a result of mining activity.

Section III The Model with Discrimination of Predictors:

\[
Y = 12.3 + 0.241 X_5 + 0.527 X_13 \pm \epsilon
\]

Which has a R² of 89.09%, while the complete model with 14 predictor variables has a R² of 90.10%, has lost a 0.01% confidence in the predictions, this indicates us with more number predictor variables can be modeled with greater approximation to reality. It presents a contamination index of 0.50 ie medium contamination, in the trajectory it recovers naturally. Quality rating of 42.00 representing inadequate (contaminated) quality.

The zoning of the Water Quality Index during the course of the Suches River, using ICASUCHES, allowed us to establish that, in general, during the last years (2009-2014) and the three climatic conditions, the river presents an acceptable quality in Section I of Laguna Suches - Hito Nro 19), an inadequate quality in Section II (Hito No. 19 - Puente...
Castilla), in which the greatest deterioration of the water quality of the Suches river is reached, and finally a slight recovery of the quality in Section III (Puente Castilla - Lake Titicaca mouth), despite being still classified as of inadequate quality.

References

the pipeline Cusiana-Coveñas, Cali, Colombia: Pollutant Control Research Center. University of the Valley.


