# Application of multivariate methods and geoestatistics to model the relationship between $CO_2$ emissions and physicochemical variables in the Hidrosogamoso reservoir, Colombia

Aplicação de métodos multivariados e geoestatísticos para modelar
a relação entre as emissões de $CO_2$ e as variáveis fisicoquímicas no
reservatório de Hidrosogamoso, Colômbia

Ingry Natalia Gómez Miranda[1]* ⓘ, Fabio Vélez Macías[2] ⓘ and Gustavo Antonio Peñuela Mesa[1] ⓘ

[1]Group of Diagnostic and Control of the Pollution - GDCON, Environmental School, Faculty of Engineering, University of Antioquia UdeA; Street 70 No. 52-21, Medellín, Colombia

[2]Group of Investigation in Geography, Limnology, and Environmental Modelation – GEOLIMNA, Environmental School, Faculty of Engineering, University of Antioquia UdeA; Street 70 No. 52-21, Medellín, Colombia

*e-mail: ingry.gomez@udea.edu.co

**Abstract: Aim:** This article deals with the estimation of a model for $CO_2$ emissions in the Hidrosogamoso reservoir based on the organic matter level and water quality. This is in order to determine the impact of the creation of a tropical reservoir on the generation of greenhouse gases (GHG), and to establish the water quality and emissions dynamics. We hypothesize that the spatial variability of emissions is determined by water quality and carbon cycling in water. **Methods:** Multivariate techniques were applied to determine the relationships between $CO_2$ and certain physicochemical variables measured in the reservoir between February and May 2015, taking samples in 10 stations and measuring 14 variables (water quality parameters and $CO_2$). Factor, cluster, discriminant and regression analysis, as well as the geostatistical technique kriging, were used. **Results:** We observed that all variables except dissolved organic carbon have strong linear relationships. Nitrate, total-P, total solids and total suspended solids are related due to the presence of nutrients in the water; chlorophyll *a* and biodegradable dissolved organic carbon due to organic carbon; and alkalinity and dissolved solids due to dissolved minerals. The sampling stations can be classified into two homogeneous groups. The first consists of the stations peripheral to the reservoir and the second of stations inside the reservoir. This difference is due mainly to the behavior of chlorophyll *a* and biodegradable dissolved organic carbon, and these two variables are also the best predictors for $CO_2$, with a maximum adjustment of 70%. **Conclusions:** Our main conclusion is that the production of $CO_2$ is due to decomposition of flooded organic carbon, depends on the soils flooded and the tributary water quality, and that the production of this gas will, based on the literature, continue for 5 to 10 years depending on the nature of the forest flooded.

**Keywords:** $CO_2$ emissions; hydropower; organic carbon; tropical reservoir; water quality.

**Resumo: Objetivo**: Este artigo trata de estimar um modelo de emissões de $CO_2$ no reservatório Hidrosogamoso a partir da matéria orgânica e da qualidade da água, para determinar o impacto da criação de um reservatório tropical na geração de gases de efeito estufa (GEE) e estabelecer a dinâmica da qualidade da água e das emissões. Nossa hipótese é que a variabilidade espacial das emissões é forçada pela qualidade da água e pelo ciclo do carbono na água. **Métodos:** Técnicas multivariadas foram aplicadas para determinar as relações entre o $CO_2$ e determinadas variáveis físico-químicas medidas no reservatório entre fevereiro e maio de 2015, tomando amostras em 10 estações e medindo 14 variáveis

(parâmetros de qualidade da água e $CO_2$). Foram usadas as técnicas estatísticas de Fator, Cluster, Análise Discriminante e Regressiva, bem como a técnica geoestatística de krigagem. **Resultados:** Observamos que todas as variáveis, exceto o carbono orgânico dissolvido, possuem fortes relações lineares. Nitrato, P-total, sólidos totais e sólidos suspensos totais estão relacionados devido à presença de nutrientes na água; clorofila *a* e carbono orgânico dissolvido biodegradável devido ao carbono orgânico; e alcalinidade e sólidos dissolvidos devido a minerais dissolvidos. As estações de amostragem podem ser classificadas em dois grupos homogêneos. O primeiro consiste nas estações periféricas do reservatório e a segunda das estações no interior do reservatório. Essa diferença é devido principalmente ao comportamento da clorofila e do carbono orgânico dissolvido biodegradável, e essas duas variáveis também são os melhores preditores para o $CO_2$, com um ajuste máximo de 70%. **Conclusões:** Nossa principal conclusão é que a produção de $CO_2$ é devido à decomposição do carbono orgânico inundado, aos solos inundados e à qualidade da água dos afluentes, e que a produção deste gás continuará, de acordo com a literatura, por 5 ou 10 anos, dependendo da natureza da floresta inundada.

**Palavras-chave:** emissões de $CO_2$; hidrelétrica; carbono orgânico; reservatório tropical; qualidade da água.

## 1. Introduction

Hydroelectricity constitutes roughly 16% of the world's electrical generation according to the 2015 Hydropower Status Report (IHA, 2015) of the International Hydropower Association. In 2015, 33.7 GW of new capacity was installed, including 2.5 GW of pumped storage, bringing total hydropower capacity to 1,212 GW worldwide (IHA, 2016). This form of energy production has been considered a clean source of energy for a long time. However, it has recently been shown that some reservoirs emit greenhouse gases (GHG) at levels equivalent to those generated by fossil fuels for comparable quantities of energy (Fearnside, 2004), (Fearnside, 2016a).

The main GHG are carbon dioxide ($CO_2$), methane ($CH_4$) and nitrous oxide ($N_2O$) (IPCC, 2001). These gases are emitted into the atmosphere from natural aquatic sources, terrestrial ecosystems and anthropogenic sources (Tremblay, et al., 2011). Since the end of the pre-industrial era in the 1750s, concentrations of these GHG have increased by 43% ($CO_2$), 152% ($CH_4$) and 20% ($N_2O$) (IPCC, 2007). $CO_2$ is extremely soluble in water, some 200 times more so than oxygen (Wetzel & Likens, 2013). It is the largest contributor to the emission of GHG and has accounted for approximately 77% of global warming potential (GWP) weighted emissions since 1990, making up 75% of total GWP-weighted emissions in 1990 and 77% in 2013 (USEPA, 2015).

In tropical reservoirs, various studies have been carried out in relation to the generation of GHG (Fearnside, 1995, 2013, 2015a, 2015b, 2016b; Galeotti & Lanza, 1999; Galy-Lacaux et al., 1999; Delmas et al., 2001; Rosa et al., 2004; Abril et al., 2005; Guérin et al., 2006; Rosa et al., 2006; Gunkel, 2009; Tremblay et al., 2011; Barros et al., 2011; Kemenes et al. 2011, 2016; De Faria et al.,

2010). $CO_2$ has a significant influence on the dynamics of gaseous emissions in reservoirs. This is especially so in the tropics, because of the large quantities of organic carbon subject to flooding in the rainy regions of tropical forests, in addition to high temperatures that favour the decomposition of carbon (Delmas et al., 2001). The amounts of GHG emitted vary greatly depending on the geographical location, the age of the reservoir, external inputs of carbon and nutrients, and characteristics of the reservoir such as water flow, turnover time, area, depth, water level fluctuations and the positioning of the turbines and spillways (Fearnside, 2016b).

The importance of precise GHG emission patterns for a predictive understanding of GHG budget across temporal and spatial scales has been highlighted (Zhao et al., 2013). Some recent studies focus on the understanding of these dynamics (Barros et al., 2011; Diem et al., 2012; De Faria et al., 2010; Kemenes et al., 2016).

Colombia, like other tropical countries, has huge hydroelectric potential, and is the 3rd highest producer of hydroelectric energy in South America, with an installed capacity of 10,793 MW. Of these, 820 MW are produced by Hidrosogamoso, representing approximately 8.3% of the electrical energy supply for the country.

Considering the above, we hypothesize that spatial variability of emissions is determined by water quality and carbon cycling in water. The aim of this study was to model $CO_2$ emissions in the Hidrosogamoso reservoir based on organic matter and water quality, in order to determine the impact of the creation of a tropical reservoir on the generation of GHG and to establish the dynamics of the water quality and emissions. The specific objectives of the research were: 1) to establish the relationship between water quality and organic carbon in order to find underlying factors and

estimate latent variables to evaluate their impact on $CO_2$ generation; 2) to determine spatial dynamics through analysis of state variables of emissions and water quality in the Hidrosogamoso reservoir system using grouping methods and spatial modeling for the identification of critical zones in the production of GHG; 3) to describe the relationship between $CO_2$ generation; water quality and organic carbon in order to establish the significant variables of GHG generation. The study aims to contribute to the management of water resources in Hidrosogamoso and to the determination of the impacts of future hydropower projects on regional and global carbon balance, especially in tropical areas where most of the projected new dams are located.

## 2. Methods

### 2.1 Study area

Hidrosogamoso (Figure 1) is located in Santander, Colombia (7° 6'11.57"N, 73°24'37.09"W) in an area where the Sogamoso River crosses the Serranía de la Paz, 75 km from the mouth of the River Magdalena and 62 km from the confluence of the Suárez and Chicamocha rivers. The primary part of the Sogamoso River used in the generation of electrical energy features a 190 m high dam and a powerhouse of subterranean machines with the three biggest generators in Colombia. With 820 MW of installed capacity and an annual average generation of 5.056 GWh per year, it represents the fourth largest installed capacity in the country. Table 1 presents the main characteristics of the reservoir.

### 2.2 Physical Characteristics

The zone in which the reservoir is inserted is characterized by rains caused by movements of the intertropical confluence zone during the year. In the first half of the year, this zone displaces from south to north and produces an increase in precipitation for the months of April and May. In the second part of the year a movement of the zone from north to south generates rains in the months of October and November. Likewise, in the second semester the precipitation is higher because the movement of the intertropical confluence zone brings masses of air saturated with water from the Atlantic Ocean.

In the reservoir area, two centres of high rainfall can be observed: one over the Opón Swamp and another between the River Sogamoso and *la ciénaga de Paredes*. The high level of sunlight in the region is due to the small number of of geographic obstacles. This allows the land to receive a greater number of

**Table 1.** Discharge average (Discharge), Area of the tributary basin (Area), Water mirror surface area (Surface), Total Volume (Volume) of the Hidrosogamoso reservoir.

| Characteristic | Value |
|---|---|
| Discharge average | 474.6 m³.s⁻¹ |
| Area of the tributary basin | 2,117,800 ha |
| Water mirror surface area | 6,960 ha |
| Total volume | 4,800 Mm³ |
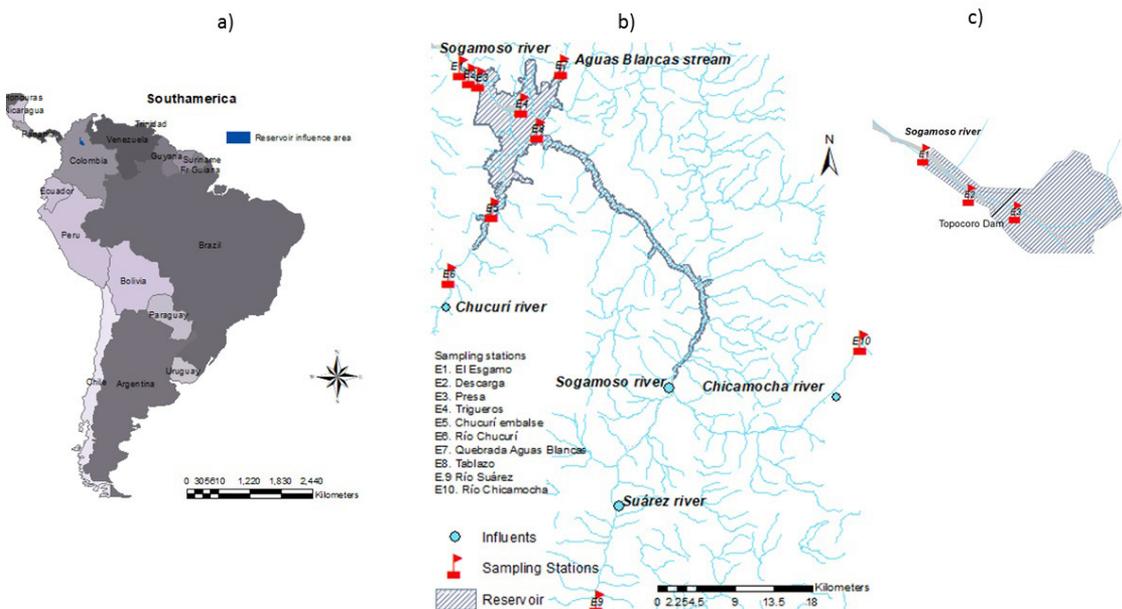


**Figure 1.** Study area: a) Colombia in South America; b) Hidrosogamoso reservoir an sampling stations; c) Hidrosogamoso dam (Topocoro).

hours of sunlight during the day, unlike regions with broken reliefs. This phenomenon, in combination with high temperatures of 27 °C and above, determines the high levels of evaporation that favour connective-type precipitation. The humidity average is greater than 80%, which provides a continuous availability of water on the ground throughout the year (GIDROT et al., 2011).

## 2.3 Hydrological Characteristics

The reservoir's major tributaries are the Chicamocha, Suárez, Sogamoso and Chucurí rivers and the Aguas Blancas stream. Table 2 shows their main characteristics.

## 2.4 Variables monitored

The reservoir monitoring begun before impoundment and continued after it. For this study we used samples taken in 2015, the first year after impoundment. The sample size was n=20, with 2 sampling campaigns in 10 sampling stations. We measured the following variables: total alkalinity, total phosphorus, dissolved solids, total suspended solids, total solids, dissolved organic carbon, biodegradable dissolved organic carbon, chlorophyll "*a*", biological oxygen demand after 5 days, chemical oxygen demand, ammonia nitrogen, total Kjeldahl nitrogen, nitrates and carbon dioxide. The samples were collected following the criteria established by the Standard Methods (APHA, 2012) and the sampling protocols of the laboratory of the "Pollution Control and Diagnosis Group-GDCON". After collecting the samples of $CO_2$ in the Tedlar stock exchanges, we transported them at environmental temperature to the laboratory (via land route to avoid abrupt

changes of pressure), where the gases were later analysed using a 7890To gas chromatograph with a 5975C mass spectrometer and a conductivity detector. The parameters monitored, abbreviations used, units and reference methods are shown in Table 3.

## 2.5 Data processing, multivariate statistical and geostatistical methods

The variables $DBO_5$, DQO, ammonia-N and NTK were discarded because their measurements showed values below their respective quantification limits of 5 mg $DBO_5$.$L^{-1}$, 25 mg $O_2$.$L^{-1}$, 5 mg$NH_3^-$N.$L^{-1}$ and 5 mgN.$L^{-1}$.

The Kolmogorov - Smirnov (KS) nonparametric test was performed to determine whether the data followed a normal distribution. According to the test, any variable from a normal distribution with a 95% confidence level cannot be rejected. The correlations between the variables (except $CO_2$ as the endogenous variable in the regression) were calculated using the Pearson correlation coefficient and organized in the correlations matrix. A correlation is considered strong (correlation is bounded between -1 and 1) if the coefficient is greater than 0.5 or lower than -0.5. The factor analysis multivariate statistical method was then used to reduce the dimensionality of the data (high correlations) and find underlying factors that generate variables (Thurstone, 1931). Heirachical cluster analysis was used to identify potential clusters in the sampling stations (Rencher & Christensen, 2003). Discriminant analysis was applied to confirm the clusters and to determine the variables that grouped these statistics (Johnson & Wichern, 2007).

**Table 2.** Hidrosogamoso major tributaries main characteristics: Formation, River mouth, Multiannual average discharge (Discharge) and Basin size (Basin).

| Influent | Formation | River mouth | Multiannual average discharge | Basin size |
|---|---|---|---|---|
| Chicamocha River | Originating in the province of Boyacá | Sogamoso River | 135 m³.s⁻¹ | 1,633,200 ha |
| Suárez River | Originating in the lagoon of Fúquene | Sogamoso River | 195 m³.s⁻¹ | 982,300 ha |
| Sogamoso River | Originating in the confluence of the Chicamocha and Suárez Rivers | Magdalena River | 540 m³.s⁻¹ | 111,944 ha |
| Chucurí River | Originating in the high part of San Vicente de Chucurí | Sogamoso River | 8.7 m³.s⁻¹ | |
| Aguas Blancas stream | Originating in the department of Santander | Sogamoso River | 0.88 m³.s⁻¹ | |

**Table 3.** Water quality (physicochemical) and Greenhouse gases variables, units and methods of reference measured to the model estimation.

| Parameter | Abbreviation | Units | Reference method |
|---|---|---|---|
| Total Alkalinity | Alkalinity | $mgCaCO3.L^{-1}$ | 2320 B* |
| Total Phosphorus | Total-P | $mgP.L^{-1}$ | 4500 P E* |
| Dissolved Solids | DS | $mg.L^{-1}$ | 2540 C* |
| Total Suspended Solids | TSS | $mg.L^{-1}$ | 2540 D* |
| Total Solids | TS | $mg.L^{-1}$ | 2540 B* |
| Dissolved Organic Carbon | DOC | $mg\ C.L^{-1}$ | 5310 B* |
| Biodegradable Dissolved Organic Carbon | BDOC | $mg\ C.L^{-1}$ | Internal method |
| Chlorophyll a | Chlorophyll | $mg.L^{-1}$ | 10200 H* |
| Biological Oxygen Demand at 5 days | BOD5 | $mg\ DBO_5.L^{-1}$ | 5210 D* |
| Chemical Oxygen Demand | COD | $mg\ O_2.L^{-1}$ | 5220 D* |
| Ammonia Nitrogen | Amonia-N | $mg\ NH_3N.L^{-1}$ | 4500 NH3 B C* |
| Total Kjeldahl Nitrogen | TKN | $mg\ N.L^{-1}$ | 4500 Norg B* |
| Nitrates | Nitrates | $mg\ NO_3^-.L^{-1}$ | 4110 B* |
| Carbon Dioxide | $CO_2$ | $g\ CO_2.m^{-2}.d^{-1}$ | Internal method |

*(APHA, 2012).

Multiple regression was used to determine a model that explained the existent relation between the variables analyzed and $CO_2$ at the post-filling stage of the Hidrosogamoso reservoir. Finally, spatial changes in the $CO_2$ and physicochemical variables were analysed with the geostatistical interpolation technique kriging. All mathematical and statistical calculations were performed using Microsoft Office Excel 2010 (SN: RJRGF-8H8PH-R4B8M-WBMBQ-9TDH9), Statgraphics Centurion XVI (SN: S520-6BOA-56S1-YKON-3EM1) and Stata 13 (1990516033). Geostatistical analysis and construction of the maps were done in ArcMap 10.2.2 (ID: 4401505451).

### 2.5.1 Factor Analysis

Factor analysis (FA) is intended to explain a set of variables observed by a small number of latent or unobserved variables called factors (Peña, 2002). In factor analysis the variables are represented as linear combinations of a small set of random variables (with m <<p) called factors. Where the factors are *latent constructs* that generate the variables. The factor model is:

$$(\mathbf{X} - \boldsymbol{\mu}) = \mathbf{LF} + \boldsymbol{\varepsilon} \tag{1}$$

where $\boldsymbol{\mu}$ is the vector of averages of $\mathbf{X}$, $\mathbf{L}$ is the matrix that contains the loads or weights of the jth factor $f_j$ is the ith variable $X_i$, $\mathbf{F}$ is the matrix that contains the factors and $\boldsymbol{\varepsilon}$ is the vector of errors that identifies the part of the variable that is divergent from the other variables.

### 2.5.2 Cluster analysis

Cluster Analysis (CA) aims to find an optimal grouping in which the observations or objects within each cluster (group) are similar but the clusters are different from each other (Rencher & Christensen, 2003). There are two approaches to cluster analysis: hierarchical and non-hierarchical. The differences between the two methods are detailed in Johnson (1998) and depend on the aim of the research. The hierarchical clustering approach is more common because the groupings are observable on a graph called a dendrogram, making it possible to visualize the formation of groups.

Cluster analysis usually involves at least three steps. The first is the measurement of some form of similarity or association between the entities (observations or objects), to determine how many groups really exist in the sample. The Euclidean or square Euclidean distance is usually used for this. The second step is the actual clustering process, whereby entities are partitioned into groups (clusters). Optimal results can be obtained using Ward's method. The final step is to profile the entities to determine their composition. This profiling is often achieved by applying discriminant analysis to the groups identified by the cluster technique (Hair et al., 2010).

### 2.5.3 Discriminant analysis

Discriminant analysis (DA) is a multivariate technique used to determine the variables responsible for the separation of the units within groups (Bierman et al., 2011). DA can resolve

a series of questions including, but not limited to, the determination of statistical differences between two or more groups. This determines which independent variables contribute to the majority of the differences between the groups and makes it possible to establish procedures to classify units inside the groups (Hair et al., 2010). DA can be considered as regression analysis where the endogenous variable *Y* is categorical, taking values or categories for each group, and the exogenous variables are the continuous variables that determine to which group the units belong.

2.5.4 Regression analysis

The models used in regression analysis provide equations that describe the relationships of dependence between one or more endogenous variables (denoted as Y's) explained by a set of exogenous variables (denoted as X's). This results in the model shown below:

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} \tag{2}$$

where **Y** is the matrix containing the endogenous variables (explained), **X** is the matrix of exogenous variables (explanatory), **β** contains the weights or marginal contributions that have X's in Y's, and **e** is the error vector.

2.5.5. Geostatistical methods of interpolation: kriging

Geostatistical methods are widely used in exploration and reservoir modelling. They have the advantage of quantifying and predicting spatial variability in order to build models and make decisions (Chihi et al., 2013). Giraldo (2002) states that kriging is a method of interpolation from geostatistics that predicts values of a variable at unsampled locations based on the information

collected. The interpolation process is performed as follows: Measurements of the variable of interest Z are made in points of the study area. In this case, realizations of the variables are obtained and it is necessary to predict in the when there are no measurements. In these circumstances, the ordinary kriging method proposes that the value of the variable can be predicted as a linear combination of the n random variables as follows:

$$Z^*_{x_0} = \lambda_1 Z(x_1), \lambda_2 Z(x_2), \ldots, \lambda_n Z(x_n) = \sum_{i=1}^{n} \lambda_i Z(x_i) \tag{3}$$

where $\lambda_i$ represents weights of the original values. These weights are calculated based on the distance between the sampled points and the point where the corresponding prediction will be made. The sum of the weights must be equal to one because the expected value of the prediction must be equal to the expected value of the variable. This is known as the requirement of unbiasedness.

## 3. Results and Discussion

### 3.1. Statistical analysis

3.1.1. Factor analysis

The correlations between physicochemical variables exogenous variables are shown in Table 4. Here it is observed that all variables except DOC have strong linear relationships. For example, the strongest relationship is 0.93 for the total-P and TSS, the explanation for which is that more than 90% of the phosphorus in fresh water bodies (like reservoirs) exists in the form of organic phosphates and cellular constituents in the biota and is adsorbed to inorganic and dead particulate organic materials (Wetzel, 2001) as dissolved solids. The maximum inverse relationship occurs between BDOC and nitrates (-0.72). According to Wetzel (2001), a

**Table 4.** Correlation matrix of the physicochemical variables studied in Hidrosogamoso during the first year after impoundment. Variables: Nitrates, Alkalinity, Total Phosphorous (P_Total), Dissolved Solids (DS), Total Suspended Solids (TSS), Total Solids (TS), Dissolved Organic Carbon (DOC), Biodegradable Dissolved Organic Carbon (BDOC) and Chlorophyll.

| | Nitrates | Alkalinity | Total-P | DS | TSS | TS | DOC | BDOC | Chlorophyll |
|---|---|---|---|---|---|---|---|---|---|
| Nitrates | | | | | | | | | |
| Alkalinity | -0.1044 | | | | | | | | |
| P_Total | 0.797 | -0.775 | | | | | | | |
| DS | 0.0678 | 0.7166 | -0.081 | | | | | | |
| TSS | 0.6506 | -0.4852 | 0.9312 | -0.37 | | | | | |
| TS | 0.5845 | -0.2524 | 0.485 | 0.2125 | 0.457 | | | | |
| DOC | -0.2007 | -0.456 | 0.0259 | 0.087 | 0.1887 | 0.467 | | | |
| BDOC | -0.7199 | -0.026 | -0.3855 | 0.122 | -0.2542 | -0.3413 | 0.4091 | | |
| Chlorophyll | -0.5977 | 0.3077 | -0.55 | 0.3123 | -0.5538 | -0.3442 | -0.0941 | 0.7426 | |

decrease in nitrogen occurs with increasing organic carbon content in lakes, because water bodies containing a high concentration of dissolved organic matter receive increasingly greater proportions of their organic content from organic plant material produced in wetland and littoral marsh areas surrounding the water. The inverse relationship between TSS and chlorophyll is explained because the high presence of suspended solids blocks light input, which is essential for photosynthesis, thereby interrupting the production of chlorophyll. The inverse relationship between total-P and chlorophyll is explained by the consumption of nutrients by photosynthetic organisms. Primary producers (chlorophyll producers) are potentially limited by carbon, nitrogen or phosphorus. The last of these, followed by nitrogen, is most likely to limit algal production in lakes, where it is less abundant in solution (Hairston & Fussmann, 2001).

The dependence between water quality (alkalinity, phosphorus and nitrates) and organic carbon (dissolved solids, total suspended solids, total solids, biodegradable dissolved organic carbon and chlorophyll) can be explained by the anthropic interventions that occur in the area along the watershed and by the variation in precipitation and flow during the study period (Coletti et al., 2010). Some investigations, Coletti et al. (2010); Shrestha & Kazama, (2007); Sun et al. (2016); Wu et al. (2014), report that wastewater discharges and the nutrients that reach the reservoirs are the main factors responsible for changes in water quality. These changes generate favorable conditions for GHG emissions not only in the reservoir, but in the rivers themselves (Demarty & Bastien, 2011; Rasera et al., 2008; Sawakuchi et al., 2014; Zhao et al., 2013).

Due to these high correlations, DOC is excluded in the factor analysis due to its low coefficient. Therefore, we can say that there are underlying factors or latent variables that generate data (Rencher & Christensen, 2003). Accordingly, factor analysis was applied in order to estimate these latent variables using maximum likelihood, due to the normality of the variables. Varimax rotation was used to make the variables orthogonal and so permit their use in subsequent analyses. The result is three factors with a cumulative variance percentage of 100% without loss of information, as shown in Table 5.

The factor loadings in each variable are shown in Table 6, where the most weighted factor determines the largest contribution of a factor to a variable,

**Table 5.** Factor Analysis for the physicochemical variables measured: three factors have been extracted, because they had an eigenvalue greater than 1.0. It accounts for 100% of the variability in the original data.

| Factor | Factor eigenvalue | Total variance percentage | Total cumulative variance percentage |
|--------|-------------------|---------------------------|--------------------------------------|
| 1 | 2.647 | 40.9% | 40.9% |
| 2 | 2.016 | 31.1% | 72% |
| 3 | 1.813 | 28% | 100% |

**Table 6.** Factor loadings for the physicochemical variables: the equations estimate the latent factors after rotation has been performed. Variables: Nitrates, Alkalinity, Total Phosphorous (P_Total), Dissolved Solids (DS), Total Suspended Solids (TSS), Total Solids (TS), Biodegradable Dissolved Organic Carbon (BDOC) and Chlorophyll.

| Variable | Factor 1 | Factor 2 | Factor 3 |
|----------|----------|----------|----------|
| Nitrates | 0.666 | -0.6263 | 0.1155 |
| Alkalinity | -0.1904 | -0.0529 | 0.8918 |
| P_Total | 0.9703 | -0.242 | -0.0062 |
| DS | -0.0483 | 0.1184 | 0.8827 |
| TS | 0.4296 | -0.28 | -0.0725 |
| TSS | 0.9285 | -0.1165 | -0.3527 |
| Chlorophyll | -0.3923 | 0.6922 | 0.309 |
| BDOC | -0.1508 | 0.9886 | -0.0027 |

regardless of the sign. This is how Factor 1 generates Nitrates, total-P, ST and SST and, as explained in the correlations, the reason for this is that a large percentage of nutrient loads are often associated with suspended particles (Thornton et al., 1990). Factor 2 generates Chlorophyll and BDOC, as chlorophyll is a biodegradable organic compound which generates dissolved carbon in the water. Factor 3 generates alkalinity and SD, through minerals dissolved in the water.

### 3.1.2 Cluster analysis and discriminant analysis

In order to find similarities in the sampling stations, we applied cluster analysis using the squared Euclidean distance and Ward's method. The variables Factor 1, Factor 2, Factor 3, DOC and $CO_2$ were used. Figure 2 shows the resulting dendrogram, where it is clear that the stations are divided into two groups: Cluster 1, comprised by E1 (*Esgamo*), E10 (*Chicamocha River*), E6 *(Chucurí River)* and E9 *(Suarez River)*; and Cluster 2, comprised by E7 *(Aguas Blancas stream)*, E5 *(Chucurí reservoir)*, E8 *(Tablazo)*, E4 *(Trigueros)*, E2 *(Outlet)* and E3 *(Dam)*. These clearly show that each group can be defined by its spatial location,

as shown in Figure 1. Cluster 1 contains stations on the periphery of the reservoir, while Cluster 2 consists of those located inside it. Although the information reported covers a short period of time, it can be said that the spatial parameters and $CO_2$ behave differently in the water of the periphery compared to that within the reservoir. In Shrestha & Kazama (2007), it was found through hierarchical cluster analysis that the number of clusters was also determined by the spatial characteristics of the site of study. Similary, in Varol et al. (2012), each cluster represented the geographical location of sampling sites located at each reservoir dam. Another cluster analysis is presented in Belkhiri & Narany (2015), where two clusters defined by the changes in the

physicochemical characterictics of the water were formed. This confirms the influence of spatial distribution on water quality.

To monitor the water analysis we used DA for as a complement to CA, as in Bierman et al. (2011); Shrestha & Kazama (2007) and Varol et al. (2012). It was found that the cluster yielded a correct classification in 100% of cases, as shown in Table 7. The classification functions are shown in Table 8, where the determining variables in the standings are $CO_2$ and Factor 2 (composed of Chlorophyll and BDOC) which means that these three variables are the main determinants of spatial variability due to their high $R^2$ value and p significance below 0.05.
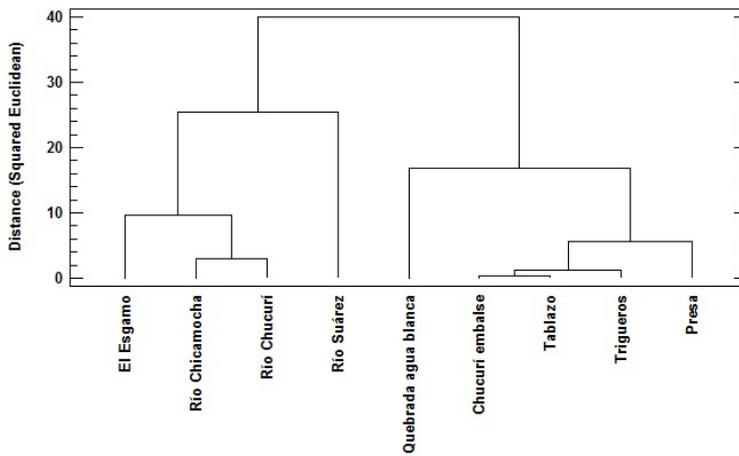


**Figure 2.** The dendogram shows the result of clustering the n = 9 sampling stations in the Hidrosogamoso reservoir using Ward's method and squared Euclidian distance. El Esgamo, Río Chicamocha, Río Chucurí and Río Suárez stations forms a first cluster separate from the second one formed by Quebrada agua blanca, Chucurí embalse, Tablazo, Trigueros and Presa stations. The firts one located at the peripheria and the second one near the center of the reservoir.

**Table 7.** Classification Table for the two groups from cluster analysis. Each row tabulates the results for cases that actually belong to a particular group and the columns show how often they were classified as belonging to each group. The percentage of cases correctly classified is 100%.

| Actual Cluster | Group Size | Predicted Cluster | |
|---|---|---|---|
| | | 1 | 2 |
| 1 | 4 | 4 (100.00%) | 0 (0.00%) |
| 2 | 5 | 0 (0.00%) | 5 (100.00%) |

Percentage of cases correctly classified: 100.00%.

**Table 8.** Classification Function Coefficients and Analysis of Variance. Variables: Factor 1 (Nitrates, Total-P, ST and SST), Factor 2 (Chlorophyll and BDOC), Factor 3 (Alkalinity and SD) and $CO_2$. F ratio= F-test for ANOVA, Significance with 95% confidence level = p-value, determination coefficient =$R^2$.

| Variable | Function 1 | Function 2 | F ratio | p-value | $R^2$ |
|---|---|---|---|---|---|
| DOC | 1.25675 | -1.00539 | 0.05012 | 0.8293 | 0.0071 |
| Factor 1 | 4.72025 | -3.77586 | 0.36571 | 0.5644 | 0.0497 |
| Factor 2 | -11.8049 | 9.44305 | 10.88 | 0.0131 | 0.6085 |
| Factor 3 | -13.8709 | 11.0956 | 2.3133 | 0.1721 | 0.2484 |
| $CO_2$ | -27.8051 | 22.2427 | 66.32 | 0.0001 | 0.9045 |
| CONSTANT | -23.6287 | -15.3697 | - | - | - |

The spatial variations of water quality and $CO_2$ emissions analyzed through cluster and discriminant analysis allow us to conclude that organic carbon is the main contributor to the spatial differences in emissions, due to human interventions to the ecosystem. In places where the chlorophyll and BDOC rise, $CO_2$ emissions also rise. This is shown in Figure 3, mainly in the center (around the dam in E3) where the main tributary is the Chucurí River. This zone had a "mosaic of crops, pastures, wooded pastures and natural spaces" before the impoundment, adding to which most of the urban settlements in the area discharge directly into the river, providing large amounts of organic matter that can reach the reservoir in large quantities.

The geographical location of the reservoirs has an impact on organic matter storage and water temperature, which influence GHG emissions (Yang et al., 2015). Carbon dioxide is also released from soil carbon in flooded land which, similarly to trees, is a source that will be exhausted in the future (Fearnside, 2015b). Research at the Petit Saut reservoir in French Guiana has found evidence that soil carbon is the main source for GHG produced in the initial emission pulse after the flood (Abril et al., 2005), whereas in temperate zones such as northern Canada it has been found that GHG emissions have no relation to the type of flooded soil (Duchemin et al., 1995).

### 3.1.3 Regression analysis

As reservoirs have an expected life time of over 100 years, it is essential to develop models for the generation of GHG, considering biogeochemical dynamics and spatial and temporal changes (Wang et al., 2018). In order to find an equation that describes the relationship between $CO_2$ and the physicochemical parameters, a regression model was adjusted where carbon dioxide is the endogenous variable and DOC, Factor 1, Factor 2 and Factor 3 are the exogenous variables. The resulting model, based on Equation (2) and ordinary least squares adjustment, is presented in Equation (4). Here, the only determinant for the endogenous variable was Factor 2, consisting of BDOC and Chlorophyll. The model fit was 70%.

Based on this we can conclude that BDOC and Chlorophyll are the determinants for the generation of $CO_2$ in the reservoir at the postfilling stage. This relationship is consistent with that indicated in Huttunen et al. (2002), Pérez & Restrepo (2008) and Wetzel & Likens (2013), concerning the relationship between photosynthetic activity and carbon dioxide.

The water quality of the tributaries and the area flooded with their soils, as well as the remaining vegetation, are very important for water quality. This is because they can serve as a source of nutrients for the reservoir for a long time (Gunkel, 2009). The availability of large
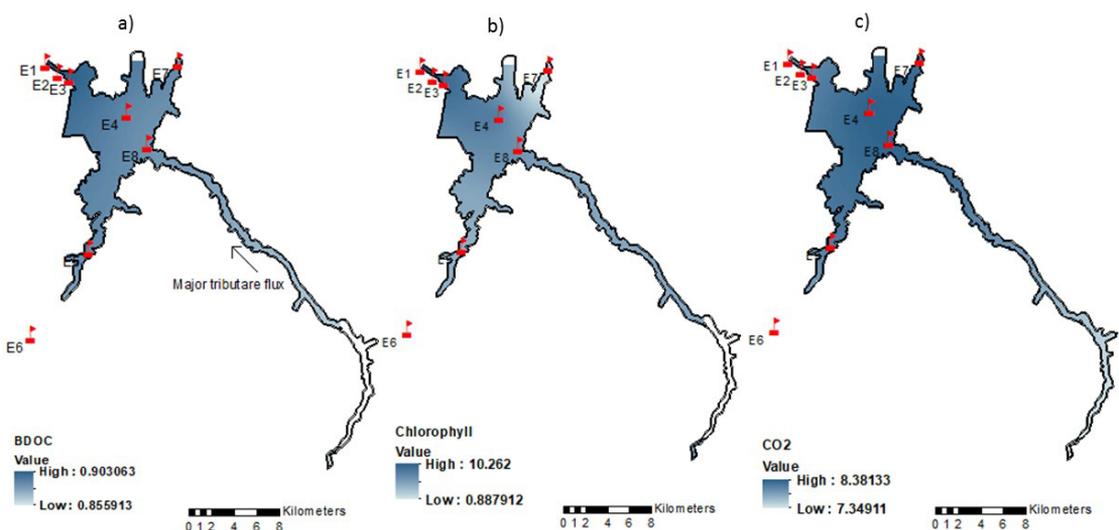


**Figure 3.** Spatial predictions using Kriging method for the distribution of a) Biodegradable Dissolved Organic Carbon (BDOC) b) Chlorophyll and c) Carbon Dioxide ($CO_2$). High levels of $CO_2$ and medium levels of chlorophyll and BDOC will be generated in areas close to the center of the reservoir (E4). Close to the dam (E1, E2, E3) will produce high values of BDOC and chlorophyll and low levels of $CO_2$. In contrast, high levels of $CO_2$ and low levels of BDOC and Chlorophyll are predicted in the major tributary area (E5, E6, E7).

amounts of dissolved organic carbon, i.e. BDOC and Chlorophyll, stimulates the generation of GHG and has been reported as being related to increases in primary productivity (Abril et al., 2005). Primary productivity has different effects with respect to the carbon balance as the reservoir ages. It may first arise from the dissolved $CO_2$ produced by decomposition of the MO, provided that the partial pressure of $CO_2$ in the water remains higher than the concentration of equilibrium with the air. In the long term, as the $CO_2$ production of the flooded CO decreases, the $CO_2$ fixed by the photosynthesis in the euphotic layer will be taken from the atmosphere, which will lead to a new flow of carbon towards the reservoir (Abril et al., 2005; Galy-Lacaux et al., 1999). Evidence has been found that in approximately 10 years these emissions decrease, as in the case of Petit Saut where emissions decreased from 200,000 tons $CO_2$.year$^{-1}$ in the first years post-flooding to 70,000 tons $CO_2$.year$^{-1}$ 10 years later.

### 3.2. Spatial distribution of $CO_2$ and Factor 2

Using the statistical analysis results described in section 3.1 we proceeded to construct maps to account for the spatial variability of these measurements, separating BDOC and Chlorophyll for better interpretation. Figure 3 presents the interpolations based on ordinary kriging.

In areas around the center of the reservoir (E8) we can see that high levels of $CO_2$ and medium levels of Chlorophyll and BDOC are predicted. In areas close to the dam (E3) we can observe high values of BDOC and chlorophyll, and low levels of $CO_2$. In contrast, high levels of $CO_2$ and low levels of BDOC and chlorophyll are predicted in the major tributary area. Those changes can be explained as follows: In a recently flooded reservoir such as Hidrosogamoso, fluxes of $CO_2$ have been reported to increase following impounding, owing to the decomposition of fresh plant biomass (Huttunen et al., 2002). Therefore, in new reservoirs, the production of $CO_2$ is mainly attributable to the decomposition of flooded organic matter and not to the phenomena of respiration and production by primary producers, the main managers of photosynthetic activity, which involve BDOC and chlorophyll. On the other hand, the high values predicted in the major tributary area are explained by the input of industrial, livestock and domestic wastewaters (Yang et al., 2015), into the rivers that feed the reservoir. Gaseous emissions are at their maximum level during the first 2 to 3 years after impounding and then slowly decrease with time (Kelly et al., 1997; Galy-Lacaux et al., 1999; Rosa et al., 2003; St. Louis et al., 2000). This shows that there is initially rapid loss of the more labile fraction of the flooded terrestrial OM, followed by a progressive decrease in quantity and bioavailability of the OM remaining at the bottom of the reservoir (Abril et al., 2005).

## 4. Conclusion

In this case study a range of statistical techniques, both multivariate and geostatistical, were applied to evaluate and model the spatial variability and relationships between various physico-chemical variables and the production of carbon dioxide in the Hidrosogamoso reservoir during the first year after impoundment.

We observed that all variables, except DOC, have strong linear relationships. Nitrates, total-P, TS and TSS are related due to the presence of nutrients in the water; chlorophyll and BDOC due to organic carbon in chlorophyll; and alkalinity and DS by dissolved minerals. The sampling stations can be classified into two homogeneous groups. The first consists of the stations peripheral to the reservoir and the second of stations within the reservoir. Based on this information, future decisions supported by cost studies could be made about reducing sampling stations. We model the relationship between carbon dioxide and measured variables, finding that chlorophyll and BDOC are the best predictors for $CO_2$, with a maximum adjustment of 70% for the period studied. Chlorophyll and BDOC are responsible for spatial variations of $CO_2$. The spatial predictions for these three variables reported that high levels of $CO_2$ and medium levels of chlorophyll and BDOC will be generated in areas close to the center of the reservoir. Meanwhile, the reservoir will produce high values of BDOC and chlorophyll and low levels of $CO_2$ in the area close to the dam. In contrast, high levels of $CO_2$ and low levels of BDOC and Chlorophyll are predicted in the major tributary area. Finally, our main conclusion is that the production of the $CO_2$ is due to the decomposition of flooded organic carbon, the production of which will continue for 5 to 10 years depending of the nature of the forest flooded, according to the literature.

The relationship between water quality and GHG generation, added to the spatial changes, allows us to conclude that there are critical zones for the generation of GHG. These are mainly in places located near the center of the reservoir, due to the nature of the flooded organic matter being

degraded, and to the high quantity of organic matter resulting from the water quality of the main tributary, which transforms into GHG. This allows us to accept the hypothesis that spatial variability of emissions is caused by water quality and carbon cycling in water.

Based on the above, the present study illustrates the usefulness of statistical and spatial analysis for the analysis and interpretation of phenomena presented in reservoirs, such as the generation of greenhouse gases.

## Acknowledgements

## References

ABRIL, G., GUÉRIN, F., RICHARD, S., DELMAS, R., GALY-LACAUX, C., GOSSE, P., TREMBLAY, A., VARFALVY, L., DOS SANTOS, M. and MATVIENKO, B. Carbon dioxide and methane emissions and the carbon budget of a 10-year old tropical reservoir (Petit Saut, French Guiana). *Global Biogeochemical Cycles*, 2005, 19(4), 1-16. http://dx.doi.org/10.1029/2005GB002457.

AMERICAN PUBLIC HEALTH ASSOCIATION – APHA. *Standard Methods for the examination of water and wastewaters*. 22th ed. Washington: American Public Health Association, 2012.

BARROS, N., COLE, J.J., TRANVIK, L.J., PRAIRIE, Y.T., BASTVIKEN, D., HUSZAR, V.L.M., DEL GIORGIO, P. and ROLAND, F. Carbon emission from hydroelectric reservoirs linked to reservoir age and latitude. *Nature Geoscience*, 2011, 4(9), 593-596. http://dx.doi.org/10.1038/ngeo1211.

BELKHIRI, L. and NARANY, T. Using multivariate statistical analysis, geostatistical techniques and structural equation modeling to identify spatial variability of groundwater quality. *Water Resources Management*, 2015, 159(3-4), 1-17. http://dx.doi.org/10.1007/s11269-015-0929-7.

BIERMAN, P., LEWIS, M., OSTENDORF, B. and TANNER, J. A review of methods for analysing spatial and temporal patterns in coastal water quality. *Ecological Indicators*, 2011, 11(1), 103-114. http://dx.doi.org/10.1016/j.ecolind.2009.11.001.

CHIHI, H., BEDIR, M. and BELAYOUNI, H. Variogram identification aided by a structural framework for improved geometric modeling of faulted reservoirs: Jeffara Basin, Southeastern Tunisia. *Natural Resources Research*, 2013, 22(2), 139-161. http://dx.doi.org/10.1007/s11053-013-9201-0.

COLETTI, C., TESTEZLAF, R., RIBEIRO, T.A., SOUZA, R.T. and PEREIRA, D.D.A. Water quality index using multivariate factorial analysis. *Revista Brasileira de Engenharia Agrícola e Ambiental*, 2010, 14(5), 517-522. http://dx.doi.org/10.1590/S1415-43662010000500009.

DE FARIA, F.A.M., JARAMILLO, P., SAWAKUCHI, H.O., RICHEY, J.E. and BARROS, N. Estimating greenhouse gas emissions from future Amazonian hydroelectric reservoirs. *Environmental Research Letters*, 2010, 10, 1-13.

DELMAS, R., GALY-LACAUX, C. and RICHARD, S. Emissions of greenhouse gases from the tropical hydroelectric reservoir of Petit Saut (French Guiana) compared with emissions from thermal alternatives. *Global Biogeochemical Cycles*, 2001, 15(4), 993-1003. http://dx.doi.org/10.1029/2000GB001330.

DEMARTY, M. and BASTIEN, J. GHG emissions from hydroelectric reservoirs in tropical and equatorial regions: Review of 20 years of CH4 emission measurements. *Energy Policy*, 2011, 39(7), 4197-4206. http://dx.doi.org/10.1016/j.enpol.2011.04.033.

DIEM, T., KOCH, S., SCHWARZENBACH, S., WEHRLI, B. and SCHUBERT, C.J. Greenhouse gas emissions (CO2, CH4, and N2O) from several perialpine and alpine hydropower reservoirs by diffusion and loss in turbines. *Aquatic Sciences*, 2012, 74(3), 619-635. http://dx.doi.org/10.1007/s00027-012-0256-5.

DUCHEMIN, E., LUCOTTE, M., CANUEL, R. and CHAMBERLAND, A. Production of the greenhouse gases CH4 and CO2, by hydroelectric reservoirs of the boreal region. *Global Biogeochemical Cycles*, 1995, 9(4), 529-540. http://dx.doi.org/10.1029/95GB02202.

FEARNSIDE, P.M. Hydroelectric dams in the Brazilian Amazon as Sources of "Greenhouse" Gases. *Environmental Conservation*, 1995, 22(1), 7-19. http://dx.doi.org/10.1017/S0376892900034020.

FEARNSIDE, P.M. Greenhouse gas emissions from hydroelectric dams: controversies provide a springboard for rethinking a supposedly "clean" energy source. *Climatic Change*, 2004, 66(1-2), 1-8. http://dx.doi.org/10.1023/B:CLIM.0000043174.02841.23.

FEARNSIDE, P.M. *Análisis de los principales proyectos hidro-energéticos en la región Amazónica*. Lima: Derecho, Ambiente y Recursos Naturales, 2013.

FEARNSIDE, P.M. *Hidrelétricas na Amazônia: impactos ambientais e sociais na tomada de decisões sobre grandes*

*obras.* Manaos: Instituto Nacional de Pesquisas da Amazônia, 2015a.

FEARNSIDE, P.M. Tropical hydropower in the clean development mechanism: Brazil's Santo Antônio Dam as an example of the need for change. *Climatic Change*, 2015b, 131(4), 575-589. http://dx.doi.org/10.1007/s10584-015-1393-3.

FEARNSIDE, P.M. Greenhouse gas emissions from Brazil's Amazonian hydroelectric dams. *Environmental Research Letters*, 2016a, 11(1), 1-3. http://dx.doi.org/10.1088/1748-9326/11/1/011002.

FEARNSIDE, P.M. Greenhouse gas emissions from hydroelectric dams in tropical forest. In: J. Lehr and J. Keeley, eds. *Alternative energy and shale gas encyclopedia*. New York: John Wiley and Sons, 2016b, pp. 428-438. http://dx.doi.org/10.1002/9781119066354.ch42.

GALEOTTI, M. and LANZA, A. Richer and cleaner? A study on carbon dioxide emissions in developing countries. *Energy Policy*, 1999, 27(10), 565-573. http://dx.doi.org/10.1016/S0301-4215(99)00047-6.

GALY-LACAUX, C., DELMAS, R., KOUADIO, G., RICHARD, S. and GOSSE, P. Long-term greenhouse gas emissions from hydroelectric reservoirs in tropical forest regions. *Global Biogeochemical Cycles*, 1999, 13(2), 503-517. http://dx.doi.org/10.1029/1998GB900015.

GIRALDO, H.R. *Introducción a la geoestadística: Teoría y Aplicación*. Bogotá: Universidad Nacional de Colombia, 2002.

GRUPO DE INVESTIGACIÓN SOBRE DESARROLLO REGIONAL Y ORDENAMIENTO TERRITORIAL – GIDROT, UNIVERSIDAD INDUSTRIAL DE SANTANDER – UIS and SECRETARÍA DE PLANEACIÓN DEPARTAMENTO DE SANTANDER – SPDS. *Santander 2030: Diagnóstico dimensión biofísico ambiental territorial de Santander*. Bucaramanga: Universidad Industrial de Santander, 2011.

GUÉRIN, F., GWENAËL, A., RICHARD, S., BURBAN, B., REYNOUARD, C., SEYLER, P. and DELMAS, R. Methane and carbon dioxide emissions from tropical reservoirs: Significance of downstream rivers. *Geophysical Research Letters*, 2006, 33(21), 1-6. http://dx.doi.org/10.1029/2006GL027929.

GUNKEL, G. Hydropower – A green energy? Tropical reservoirs and greenhouse gas emissions. *CLEAN – Soil, Air, Water*, 2009, 37(9), 726-734.

HAIR, J.F., BLACK, W.C. and BABIN, B.J. *Multivariate data analysis*. 7th ed. United States: Prentice Hall, 2010.

HAIRSTON, N.G. and FUSSMANN, G.F. *Lake ecosystems*. New York: John Wiley and Sons, Ltd., 2001.

HUTTUNEN, J.T., VÄISÄNEN, T.S., HELLSTEN, S.K., HEIKKINEN, M., NYKÄNEN, H., JUNGNER, H., NISKANEN, A., VIRTANEN, M.O., LINDQVIST, O.V., NENONEN, O.S. and MARTIKAINEN, P.J. Fluxes of CH4, CO2, and N2O in hydroelectric reservoirs Lokka and Porttipahta in the northern boreal zone in Finland. *Global Biogeochemical Cycles*, 2002, 16(1), 1-17. http://dx.doi.org/10.1029/2000GB001316.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE – IPCC. *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intergovernmental Panel on Climate Change*. New York: Cambridge University Press, 2001.

INTERGOVERNMENTAL PANEL ON CLIMATE CHANGE – IPCC. *Climate Change 2007: Synthesis Report. Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel of Climatic Change*. Geneva, Switzerland: Intergovernmental Panel on Climate Change, 2007.

INTERNATIONAL HIDROPOWER ASSOCIATION – IHA. *Hydropower Status Report 2015*. London: International Hidropower Association, 2015.

INTERNATIONAL HIDROPOWER ASSOCIATION – IHA. *Hydropower Status Report 2016*. London: International Hidropower Association, 2016.

JOHNSON, D.E. *Applied multivariate methods for data analysts*. United States: Duxbury Press, 1998.

JOHNSON, R.A. and WICHERN, D.W. *Applied multivariate statistical analysis*. United States: Pearson Prentice Hall, 2007.

KELLY, C.A., RUDD, J.W.M., BODALY, R.A., ROULET, N.P., ST. LOUIS, V.L., HEYES, A., MOORE, T.R., SCHIFF, S., ARAVENA, R., SCOTT, K.J., DYCK, B., HARRIS, R., WARNER, B. and EDWARDS, G. Increases in fluxes of greenhouse gases and methyl mercury following flooding of an experimental reservoir. *Environmental Science & Technology*, 1997, 31(5), 1334-1344. http://dx.doi.org/10.1021/es9604931.

KEMENES, A., FORSBERG, B.R. and MELACK, J.M. CO2 emissions from a tropical hydroelectric reservoir (Balbina, Brasil). *Journal of Geophysical Research*, 2011, 116(G3), 1-11. http://dx.doi.org/10.1029/2010JG001465.

KEMENES, A., FORSBERG, B.R. and MELACK, J.M. Downstream emissions of CH4 and CO2 from hydroelectric reservoirs (Tucuruí, Samuel, and Curuá-Una) in the Amazon basin. *Inland Waters*, 2016, 1(1), 1-8. http://dx.doi.org/10.1080/IW-6.3.980.

PEÑA, D. *Análisis de datos multivariantes*. España: Mac-Graw Hill Interamericana de España, 2002.

PÉREZ, G.R. and RESTREPO, J.J.R. *Fundamentos de limnología neotropical*. Medellín: Editorial Universidad de Antioquia, 2008.

RASERA, M., BALLESTER, M.V.R., KRUSCHE, A.V., SALIMON, C., MONTEBELO, L.A., ALIN, S.R.,

VICTORIA, R.L. and RICHEY, J.E. Estimating the surface area of small rivers in the southwestern Amazon and their role in CO2 outgassing. *Earth Interactions*, 2008, 12(6), 1-16. http://dx.doi.org/10.1175/2008EI257.1.

RENCHER, A.C. and CHRISTENSEN, W. F. *Methods of multivariate analysis*. New York: John Wiley and Sons, 2003.

ROSA, L.P., ELIO, M.A.U.R., SANTOS, D.O.S., MATVIENKO, B., SIKAR, E., OLIVEIRA, E. and SANTOS, D.O.S. Scientific errors in the Fearnside comments on greenhouse gas emissions (GHG) from hydroelectric dams and response to his political claiming. *Climatic Change*, 2006, 75(1-2), 91-102. http://dx.doi.org/10.1007/s10584-005-9046-6.

ROSA, L.P., SANTOS, M., MATVIENKO, B., SANTOS, E. and SIKAR, E. Greenhouse gas emissions from hydroelectric reservoirs in tropical regions. *Climatic Change*, 2004, 66(1-2), 9-21. http://dx.doi.org/10.1023/B:CLIM.0000043158.52222.ee.

ROSA, L.P., SANTOS, M.A., MATVIENKO, B., SIKAR, E., LOURENÇO, R.S.M. and MENEZES, C.F. Biogenic gas production from major Amazon reservoirs, Brazil. *Hydrological Processes*, 2003, 17(7), 1443-1450. http://dx.doi.org/10.1002/hyp.1295.

SAWAKUCHI, H.O., BASTVIKEN, D., SAWAKUCHI, A.O., KRUSCHE, A.V., BALLESTER, M.V. and RICHEY, J.E. Methane emissions from Amazonian Rivers and their contribution to the global methane budget. *Global Change Biology*, 2014, 20(9), 2829-2840. http://dx.doi.org/10.1111/gcb.12646. PMid:24890429.

SHRESTHA, S. and KAZAMA, F. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. *Environmental Modelling & Software*, 2007, 22(4), 464-475. http://dx.doi.org/10.1016/j.envsoft.2006.02.001.

ST. LOUIS, V.L., KELLY, C.A., DUCHEMIN, É., RUDD, J.W.M. and ROSENBERG, D.M. Reservoir surfaces as sources of greenhouse gases to the atmosphere: A global estimate. *Bioscience*, 2000, 50(9), 766-775. http://dx.doi.org/10.1641/0006-3568(2000)050[0766:RSASOG]2.0.CO;2.

SUN, W., XIA, C., XU, M., GUO, J. and SUN, G. Application of modified water quality indices as indicators to assess the spatial and temporal trends of water quality in the Dongjiang River. *Ecological Indicators*, 2016, 66, 306-312. http://dx.doi.org/10.1016/j.ecolind.2016.01.054.

THORNTON, K.W., KIMMEL, B.L. and PAYNE, F.E. *Reservoir limnology: ecological perspectives*. New York: John Wiley and Sons, 1990.

THURSTONE, L.L. Multiple factor analysis. *Psychological Review Company*, 1931, 38(5), 406-427.

TREMBLAY, A., VARFALVY, L., ROEHM, C. and GARNEAU, M. *Greenhouse Gas Emissions - Fluxes and Processes: Hydroelectric Reservoirs and Natural Environments*. Berlin: Springer, 2011.

U.S. ENVIRONMENTAL PROTECTION AGENCY – USEPA. *Inventory of U.S. Greenhouse Gas Emissions and Sinks: 1990-2013*. Washington: Environmental Protection Agency, 2015.

VAROL, M., GÖKOT, B., BEKLEYEN, A. and ŞEN, B. Spatial and temporal variations in surface water quality of the dam reservoirs in the Tigris River basin, Turkey. *Catena*, 2012, 92, 11-21. http://dx.doi.org/10.1016/j.catena.2011.11.013.

WANG, W., ROULET, N.T., KIM, Y., STRACHAN, I.B., DEL GIORGIO, P., PRAIRIE, Y.T. and TREMBLAY, A. Modelling CO2 emissions from water surface of a boreal hydroelectric reservoir. *The Science of the Total Environment*, 2018, 612, 392-404. http://dx.doi.org/10.1016/j.scitotenv.2017.08.203. PMid:28863371.

WETZEL, R.G. and LIKENS, G. *Limnological analyses*. New York: Springer, 2013.

WETZEL, R.G. *Limnology: lake and river ecosystems*. 3rd ed. United States: Academic Press, 2001.

WU, E.M.-Y., TSAI, C.C., CHENG, J.F, KUO, S.L. and LU, W.T. The application of water quality monitoring data in a reservoir watershed using AMOS confirmatory factor analyses. *Environmental Modeling and Assessment*, 2014, 19(4), 325-333. http://dx.doi.org/10.1007/s10666-014-9407-5.

YANG, S.-S., CHEN, I.C., LIU, C.-P., LIU, L.-Y. and CHANG, C.H. Carbon dioxide and methane emissions from Tanswei River in Northern Taiwan. *Atmospheric Pollution Research*, 2015, 6(1), 52-61. http://dx.doi.org/10.5094/APR.2015.007.

ZHAO, Y., WU, B.F. and ZENG, Y. Spatial and temporal patterns of greenhouse gas emissions from Three Gorges Reservoir of China. *Biogeosciences*, 2013, 10(2), 1219-1230. http://dx.doi.org/10.5194/bg-10-1219-2013.