Research Paper

Developing a water quality index in a tropical reservoir using a measure of multiparameters

Alberto Quevedo-Castro, Jesús G. Rangel-Peraza, Erick Bandala,
Leonel Amabilis-Sosa, Abraham Rodríguez-Mata
and Yaneth Bustos-Terrones

ABSTRACT

A water quality index (WQI) for the Adolfo López Mateos Dam (ALMD) was developed based on statistical multiparameter tools assisted with linear programming. ALMD was selected due to its social and economic significance in Sinaloa, the state with the highest agricultural production in Mexico. Twenty-six water-quality parameters were analyzed for four sampling points distributed along the dam during 2012–2017. The data were analyzed using Pearson's correlation matrix, principal components analysis (PCA) and sensitivity analysis (SA). Results indicated that variables explaining spatial and temporal water quality distribution at ALMD were total suspended solids, fecal coliforms, pH, dissolved oxygen, chemical oxygen demand, nitrate nitrogen, organic nitrogen, ammonium nitrogen, total phosphorus, orthophosphates and chlorophyll a. A series of pondering weights (*Wi*) were obtained from the PCA analysis. Every *Wi* was multiplied by the probability function of the specific parameter (*SIi*) to generate the $\mathrm{WQI}_{\mathrm{ALMD}}$ model. The model was applied to address water quality at ALMD which describes the general overall water quality in the dam as 'good'. Finally, a sensitivity analysis for the model showed that the most sensitive WQI variables were: fecal coliforms, total phosphorus, organic nitrogen, and chlorophyll a.

Key words | Pearson correlation, principal component analysis, sensitivity analysis, tropical reservoir, water quality index

Alberto Quevedo-Castro
Jesús G. Rangel-Peraza
División de Estudios de Posgrado e Investigación,
Instituto Tecnológico de Culiacán,
Juan de Dios Bátiz 310, Col. Guadalupe, P.C. 80220,
Culiacán, Sinaloa,
Mexico

Erick Bandala
Division of Hydrologic Sciences,
Desert Research Institute,
755 Flamingo Road, Las Vegas, Nevada 89119-7363,
USA

Leonel Amabilis-Sosa
Abraham Rodríguez-Mata
Yaneth Bustos-Terrones (corresponding author)
CONACYT-División de Estudios de Posgrado e Investigación,
Instituto Tecnológico de Culiacán,
Juan de Dios Bátiz 310, Col. Guadalupe, P.C. 80220,
Culiacán, Sinaloa,
Mexico
E-mail: *yabustoste@conacyt.mx*

INTRODUCTION

Reservoirs are shallow bodies of water used for population water supply by collecting, storing, and distributing water. As in other lentic water bodies, water quality in reservoirs is determined by hydrology, as well as the biological and physico-chemical parameters in the water column. These parameters, in turn, are defined by the natural and anthropogenic processes involved within the same ecosystem (Bautista & Ruiz 2011). A large amount of data is obtained in the evaluation of the water quality in a reservoir, so it is very important to summarize the data obtained in the monitoring to facilitate the communication of water characterization. Water quality indices (WQIs) are a simple tool for numerically diagnosing the quality of a water body. Water quality index (WQI) development is performed through statistical tools that allow the selection of the most representative water quality parameters in the water body and allocation of pondered weighted values for every measured parameter. WQI contributes to the management of water resources by facilitating the understanding of data generated in the evaluation of water quality parameters (Howladar *et al.* 2017).

doi: 10.2166/washdev.2018.049

There are several methodologies for WQI estimation, which vary according to the regions where they were designed (Cude 2001; Moscuzza *et al.* 2007; Tirkey *et al.* 2015). Multiparametric statistical analysis is a process that involves useful tools for decision making based on statistical procedures applied to the water quality parameters responsible for influencing the evaluation of water quality (Varol & Davraz 2015). To our knowledge, relatively few scientific reports describing water quality determination for agricultural land in subtropical regions are available and none has been developed specifically for Mexico. The aim of this work is to develop a WQI applicable to tropical reservoirs based on a methodology that relies not only on tools and multiparameter techniques. The analysis uses mathematical interactions to guarantee the developed WQI will be representative of water quality and will contribute to the monitoring work performed by regional water operators and ensure scientifically informed decision-making on water resources conservation.

MATERIAL AND METHODS

Study area

The Adolfo López Mateos Dam (ALMD) is one of the most important reservoirs for economic and social development in Mexico. ALMD provides water for the most important agricultural region of the country, and is considered a big reservoir with regard to total capacity (4,034.5 hm^3) and surface area (11,354 ha). ALMD is located 186.5 meters above sea level (m.a.s.l.) and has 1,731,400 hm^3 of total accumulated volume (CONAGUA 2015). ALMD is the reservoir with the highest capacity for annual conservation in the region, raising interest for studying such an important water body. ALMD is localized in the hydrological basin of the Culiacan River (Figure 1) and the geographical coordinates 26°03′36″–24°48′00″ N latitude and 107°16′12″–105°50′24″ W longitude.

Water quality parameters

The Mexican National Water Commission (CONAGUA) is the administrative agency responsible for management, regulation, control and protection of water in Mexico. CONAGUA monitors water quality at ALMD by determination of water quality parameters based on standard methodologies. Water quality data used in this study were obtained from CONAGUA. Average concentrations obtained for 26 water quality parameters semiannually during the period 2012–2017 in four strategic sampling sites (Figure 1) distributed throughout ALMD were used.

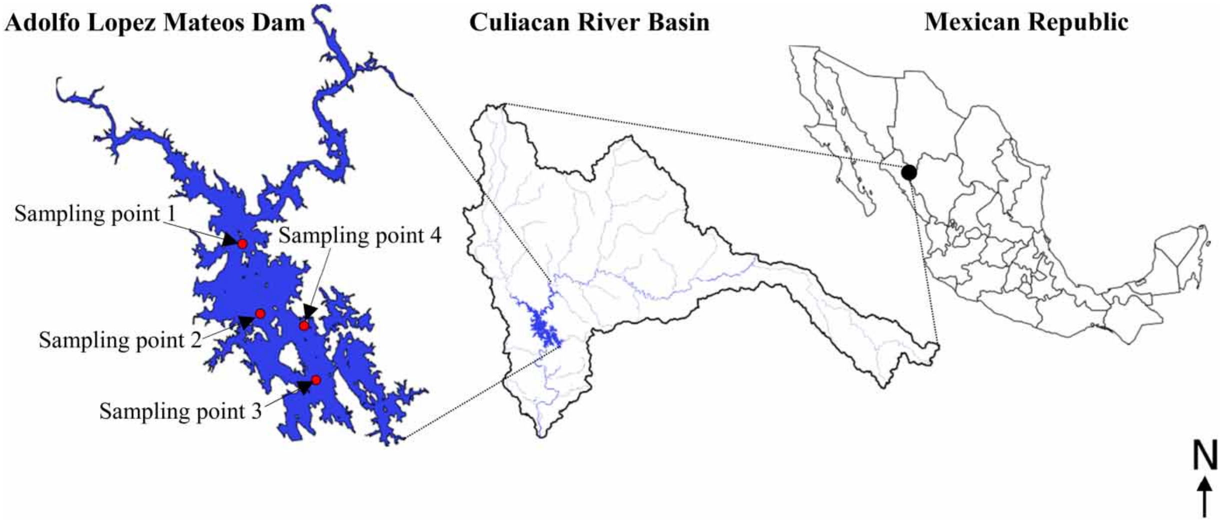

Figure 1 | Location of the Culiacan River Basin and the Adolfo Lopez Mateos Dam.

Multiparametric statistical analysis tools

The principal component analysis (PCA) was initially applied to the 26 water quality parameters set to synthesize information and generate components or strata, governed by its representativeness in analysis of variance. The analysis performed linear combinations among data in order to identify water quality parameters representing the whole set of analyzed variables without losing relevant information with statistical basis. The main components found are not correlated with each other, and each component maximizes its variance so, on many occasions, significant information is obtained from variables containing the lowest relationship between them, this being an important criterion when interpreting values and choosing the most representative study variables (An *et al.* 2015).

PCA was carried out using *Real Statistic* simultaneously with a Pearson's correlation analysis, which is an index measuring the co-variation degree between two different linearly related variables (Palácio *et al.* 2016). The Pearson's analysis was used to generate the correlation matrix and identify the amount of possible variables that would provide the same importance between variables to be discriminated against. If there is a significant correlation between variables, it means that there is redundant information, therefore, the selected parameters explain a large part of the total variability of the data.

The 26 initial water quality parameters were reduced to 11 through the application of the previously described techniques. These remaining 11 parameters were used for calculating the specific WQI for ALMD. PCA was also used, through its eigenvalues, to assign the weight of every variable used in the WQI.

WQI calculation for ALMD

Once the variable set explaining the maximum water quality variation was obtained, the following mathematical model, Equation (1), was used to estimate the ALMD water quality index (WQI_{ALMD}):

$$WQI_{ALMD} = \sum_{n=1}^{i} SI_n W_n \quad (1)$$

Table 1 | Optimal value for each SI_n

Water quality parameter	**Units**	**Optimal value ($SI_n = 100$)**
Fecal coliform	CFU/100 mL	0
Total suspended solids	mg/L	<14.14
pH	pH Units	$6.7 < pH < 7.3$
Dissolved oxygen	mg/L	≥7.7
COD	mg/L	0
Organic nitrogen	mg/L	<0.011
Ammoniacal nitrogen	mg/L	<0.011
Nitrates	mg/L	<2.5
Total phosphorus	mg/L	<0.011
Orthophosphates	mg/L	<0.09771
Chlorophyll a	mg/m^3	0.04

where SI_n is the quality function value of every water quality parameter. For every calculation obtained from SI_n, probability functions representing the distribution of every parameter with respect to its optimal value were used. In this context, SI_n value has a range of 0–100, where 100 corresponds to the best scenario of the variable (e.g. for dissolved oxygen, a value of 100 equals the saturation concentration) (Table 1).

Wi is the weight assigned to each water-quality variable based in the explained variation (% variance) of PCA results. This methodology is proposed to avoid a subjective designation of *Wi* in WQI.

To judge the WQI_{ALMD} obtained, the classification developed by the National Sanitation Foundation (NSF) was used. This classification is commonly used in other indices around the world such as the US National Sanitation Foundation Water Quality Index (NSFWQI), the Oregon Water Quality Index (OWQI), and the Canadian Council of Ministers of the Environment Water Quality Index (CCMEWQI) (Şener *et al.* 2017) because the NSF provides standardized methodologies for linear models of multiplicative summation (Brown *et al.* 1970).

Robustness of the proposed quality index and sensitivity analysis

Once each *Wi* and *SIi* was obtained, the WQI_{VALUE} was determined, which is the product of $Wi \times SIi$ for each parameter. The sum of the WQI_{VALUE} results in the WQI_{ALMD}. In order to provide statistical certainty in the procedure followed for the construction of the WQI, the values for *Wi* were varied (within the range specified in the PCA) in ranges from 0.01 and introduced in Equation (1) to obtain the different values that WQI_{ALMD} would take. The values were varied through a simplex algorithm in *MATLAB* (R2013A) to find all possible solutions of *Wi* for each variable with the restrictions derived from the PCA, so at the end all the combinations add up to 1.0.

Finally, a sensitivity analysis was performed on the 11 selected parameters to identify which are more sensitive to the proposed methodology. Sensitivity analysis consisted of fluctuating each of the water indicators (*SIi*) previously selected. This analysis was carried out using the increase and decrease of the *SIi* values of every parameter in the range between ±50% and ±30% as input variables. With regard to the output variables simulated, the results of WQI_{ALMD} display the change they had under the increase (+30% and +50%) and decrease (−30% and −50%) criteria. The parameters with greater variation are those having the greatest influence on water quality of ALMD. Once the simulation data were obtained, we calculated the sensitivity coefficient (SC), which divides the change percentage in the input variable (the result of each WQI_{VALUE} simulated) by the change percentage in the output variable WQI_{ALMD} (Equation (2)) (Rangel-Peraza 2016):

$$SC = \frac{\%\ change\ in\ WQI_{VALUE}}{\%\ change\ in\ WQI_{ALMD}} \tag{2}$$

Once Equation (2) was applied for the 11 water quality parameters, the most sensitive parameters (having greater representation in water quality spatial and temporal variation) in ALMD were identified.

RESULTS AND DISCUSSION

Data analysis

In Table 2, the results of the spatial and temporary distribution of water quality parameters in the ALMD are presented in an integrated way. Data in Table 2 also include the mean, standard error, and range for every parameter. After PCA analysis implementation, Figure 2 shows the sedimentation curve where the percentage of representativeness of every component is expressed. Sedimentation curve is composed, on the abscissa axis, of the total components resulting from PCA and, at the same time, each component is composed of a set of water quality variables. The ordinate axis represents the fraction of the total variance in water quality parameters spatial and temporal distribution. Thus, total components (26) account for 100% of the variance (Figure 2).

Table 2 | Descriptive statistical analysis

Parameter	Unit	Mean	Standard error	Range
Chlorophyll a	mg/m^3	9.2869	1.5751	37.33
Fecal coliforms	CFU/100 mL	1379	617.495	24195
Escherichia coli	CFU/100 mL	12	6.4442	245
TOC	mg/L	5.9536	0.836	33.69
BOD	mg/L	3.4918	0.3609	7.7
COD	mg/L	22.9873	1.9267	44.4
Ammoniacal nitrogen	mg/L	0.1087	0.0101	0.35
Nitrites	mg/L	0.0063	0.0017	0.06
Nitrates	mg/L	0.0675	0.0238	0.59
Organic nitrogen	mg/L	0.4364	0.0355	0.94
Total nitrogen	mg/L	0.6173	0.0389	1.03
Total Kjeldahl nitrogen	mg/L	0.5452	0.0334	0.92
Total phosphorus	mg/L	0.1127	0.0194	0.5
True color	TCU Pt/Co	18.3625	2.1632	70
Transparency	m	1.6055	0.1376	3
Conductivity	µS/cm	170.77	5.4973	130
pH	pH Units	7.9183	0.0995	2.43
Total dissolved solids	mg/L	115.4221	3.3928	74.2
Dissolved oxygen	mg/L	7.3083	0.4145	9.17
UV absorption	Abs/cm	0.0962	0.0083	0.17
Total hardness	mg/L	75.1963	5.5678	140.2
Orthophosphates	mg/L	0.0255	0.0042	0.1
Total suspended solids	mg/L	36.8493	12.9773	350
Turbidity	NTU	6.317	2.0727	73.12
Ambient temperature	°C	33.6954	0.4303	13
Water temperature	°C	29.6947	0.4435	9.3

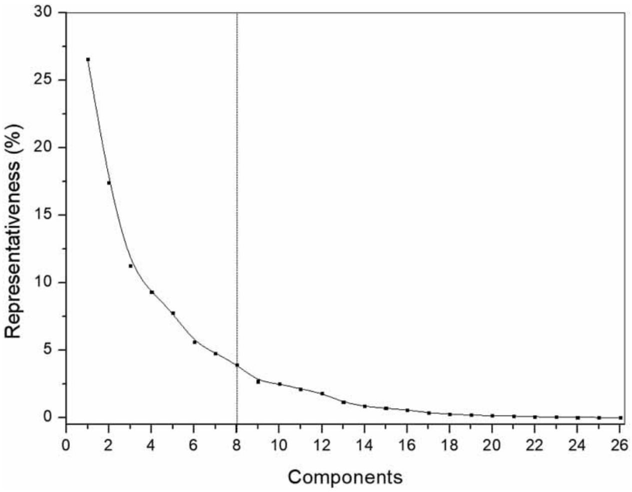

Figure 2 | Results of PCA sedimentation.

Also, Figure 2 shows that representativeness decreases considerably from component 8 and above. Therefore components 9 to 26 were discarded due to their very low representative values. Considering component 1 (PC1) to component 8 (PC8), Table 3 shows that the cumulative variance representativeness of the data is 89.34%, in agreement with various PCA methods where at least 87% cumulative variance representativeness is recommended (Ferré 1995), suggesting the reliability of the achieved results.

A Pearson correlation matrix was built (Table 4) to perform the discriminant analysis of the water quality variables with the greatest covariance. It is important considering that variables not significantly correlated are the most representative (Varol & Davraz 2015). In addition, one variable was chosen from the statistically correlated pairs of variables, based on PCA results.

Table 3 | Percentage of representativeness of the main components

Component	Eigenvalue	%	% Accumulative
1	7.1234	27.4	27.4
2	4.6135	17.74	45.14
3	3.0058	11.56	56.7
4	2.6568	10.22	66.92
5	1.8598	7.15	74.08
6	1.5529	5.97	80.05
7	1.3145	5.06	85.10
8	1.1020	4.24	89.34

Eleven representative variables were obtained as a product of statistical tools implementation (e.g., Pearson correlation and PCA), and were classified into five groups: particulate matter, nitrogen compounds, physico-chemical, phosphorus compounds and primary productivity. Once the key water quality parameters at ALMD were identified, a descriptive statistical analysis that shaped $\mathrm{WQI}_{\mathrm{ALMD}}$ construction was performed.

Principal components parameters

After the parameters with greater variance were analyzed for every component, the discriminant analysis was performed to obtain the parameters resulting from the method. Correlation between variables was considered, as well as classification by groups representing the state of water quality at ALMD. Table 5 shows the variance of parameters according to the eight main components selected, where bold numbers mean the variable with the highest covariance of each component.

Based on the analysis of variances (Table 5), the most significant parameters are in PC1, responding to variables related with solid matter (total suspended solids: 0.975 and total dissolved solids: 0.561), physico-chemical parameters (true color: 0.905 and turbidity: 0.876), microbiological variables (fecal coliforms: 0.704 and *Escherichia coli*: 0.619) and nutrients (nitrates: 0.904, nitrites: 0.766 and total phosphorus: 0.939). The parameters resulting from PC1 are essential because they respond to particulate material produced by natural or anthropogenic activity (suspended solids and dissolved solids), characteristics that impede sunlight penetration (physico-chemical), health-related pathogenic microorganisms (coliforms) and those related to primary productivity (nutrients). PC1 variables represent the largest amount of information related to water quality aspects at ALMD, being 27.4% of the data (see Table 3). In each group classification, other important variables were identified.

Particulate matter

Particulate filterable (total suspended solids >0.45 μm) and non-filterable (total dissolved solids <0.45 μm) matter is a key factor in water quality as it regulates contaminant's

Table 4 | Pearson correlation matrix

Parameter	Chlorophyll a	Fecal coliforms	*Escherichia coli*	TOC	BOD	COD	Ammonia nitrogen	Nitrites	Nitrates	Organic nitrogen	Total nitrogen	Total Kjeldahl nitrogen	Total phosphorus
Chlorophyll a	1												
Fecal coliforms	0.180	1											
Escherichia coli	−0.124	0.344	1										
TOC	0.129	0.104	−0.076	1									
BOD	0.049	−0.117	0.027	−0.192	1								
COD	−0.006	−0.037	0.003	−0.038	0.706	1							
Ammonia nitrogen	−0.153	−0.127	−0.004	−0.147	−0.010	0.375	1						
Nitrites	0.210	0.337	0.011	0.184	−0.023	0.304	0.638	1					
Nitrates	0.145	0.706	−0.023	0.223	−0.048	0.185	0.262	0.646	1				
Organic nitrogen	−0.270	−0.012	−0.213	−0.034	−0.354	−0.396	−0.346	−0.306	−0.189	1			
Total nitrogen	−0.182	0.405	−0.207	0.048	−0.350	−0.144	0.116	0.309	0.525	0.700	1		
Total Kjeldahl nitrogen	−0.329	−0.048	−0.228	−0.076	−0.378	−0.316	−0.091	−0.148	−0.128	0.966	0.775	1	
Total phosphorus	0.100	0.553	−0.066	0.227	−0.159	0.260	0.548	0.788	0.889	−0.239	0.485	−0.103	1
True color	−0.001	0.814	0.109	0.317	−0.263	0.030	0.171	0.557	0.884	−0.025	0.575	0.020	0.818
Transparency	−0.040	−0.446	0.148	−0.035	0.279	0.004	−0.120	−0.192	−0.383	−0.379	−0.620	−0.436	−0.367
Conductivity	−0.220	−0.169	−0.065	−0.210	−0.071	−0.058	0.085	−0.155	−0.136	0.531	0.419	0.587	−0.190
pH	−0.188	−0.070	−0.103	−0.347	−0.127	−0.385	−0.357	−0.338	−0.359	0.490	0.136	0.421	−0.418
Total dissolved solids	−0.250	−0.140	−0.090	−0.116	0.248	0.253	0.153	−0.089	0.089	0.292	0.355	0.352	−0.046
Dissolved oxygen	−0.321	−0.341	−0.262	−0.170	0.036	−0.014	0.009	−0.110	−0.323	0.193	−0.023	0.207	−0.235
UV absorption	−0.481	0.129	−0.028	−0.073	0.176	0.276	0.349	0.388	0.508	−0.018	0.391	0.077	0.447
Total hardness	0.206	0.026	−0.041	0.064	−0.017	0.074	0.020	−0.036	−0.027	0.136	0.111	0.150	−0.040
Orthophosphates	0.428	0.217	−0.121	−0.023	0.134	−0.006	0.154	0.506	0.492	−0.310	0.079	−0.286	0.400
Total suspended solids	0.151	0.638	−0.028	0.238	−0.084	0.283	0.486	0.835	0.901	−0.200	0.517	−0.078	0.964
Turbidity	0.179	0.842	0.012	0.183	−0.051	0.161	0.081	0.473	0.849	−0.019	0.539	0.002	0.760
Ambient temperature	−0.069	0.087	0.305	0.211	−0.217	0.060	0.148	0.115	0.079	−0.085	0.005	−0.049	0.169
Water temperature	−0.182	−0.030	0.051	0.159	−0.470	−0.177	0.242	0.149	0.063	0.450	0.510	0.544	0.144
Nitrites	0.210	0.337	0.011	0.184	−0.023	0.304	0.638	1					
Nitrates	0.145	0.706	−0.023	0.223	−0.048	0.185	0.262	0.646	1				
Organic nitrogen	−0.270	−0.012	−0.213	−0.034	−0.354	−0.396	−0.346	−0.306	−0.189	1			
Total nitrogen	−0.182	0.405	−0.207	0.048	−0.350	−0.144	0.116	0.309	0.525	0.700	1		
Total Kjeldahl nitrogen	−0.329	−0.048	−0.228	−0.076	−0.378	−0.316	−0.091	−0.148	−0.128	0.966	0.775	1	
Total phosphorus	0.100	0.553	−0.066	0.227	−0.159	0.260	0.548	0.788	0.889	−0.239	0.485	−0.103	1
True color	−0.001	0.814	0.109	0.317	−0.263	0.030	0.171	0.557	0.884	−0.025	0.575	0.020	0.818

(*continued*)

Table 4 | continued

Parameter	Chlorophyll a	Fecal coliforms	*Escherichia coli*	TOC	BOD	COD	Ammonia nitrogen	Nitrites	Nitrates	Organic nitrogen	Total nitrogen	Total Kjeldahl nitrogen	Total phosphorus
Transparency	−0.040	−0.446	0.148	−0.035	0.279	0.004	−0.120	−0.192	−0.383	−0.379	−0.620	−0.436	−0.367
Conductivity	−0.220	−0.169	−0.065	−0.210	−0.071	−0.058	0.085	−0.155	−0.136	0.531	0.419	0.587	−0.190
pH	−0.188	−0.070	−0.103	−0.347	−0.127	−0.385	−0.357	−0.338	−0.359	0.490	0.136	0.421	−0.418
Total dissolved solids	−0.250	−0.140	−0.090	−0.116	0.248	0.253	0.153	−0.089	0.089	0.292	0.355	0.352	−0.046
Dissolved oxygen	−0.321	−0.341	−0.262	−0.170	0.036	−0.014	0.009	−0.110	−0.323	0.193	−0.023	0.207	−0.235
UV absorption	−0.481	0.129	−0.028	−0.073	0.176	0.276	0.349	0.388	0.508	−0.018	0.391	0.077	0.447
Total hardness	0.206	0.026	−0.041	0.064	−0.017	0.074	0.020	−0.036	−0.027	0.136	0.111	0.150	−0.040
Orthophosphates	0.428	0.217	−0.121	−0.023	0.134	−0.006	0.154	0.506	0.492	−0.310	0.079	−0.286	0.400
Total suspended solids	0.151	0.638	−0.028	0.238	−0.084	0.283	0.486	0.835	0.901	−0.200	0.517	−0.078	0.964
Turbidity	0.179	0.842	0.012	0.183	−0.051	0.161	0.081	0.473	0.849	−0.019	0.539	0.002	0.760
Ambient temperature	−0.069	0.087	0.305	0.211	−0.217	0.060	0.148	0.115	0.079	−0.085	0.005	−0.049	0.169
Water temperature	−0.182	−0.030	0.051	0.159	−0.470	−0.177	0.242	0.149	0.063	0.450	0.510	0.544	0.144

Parameter	True color	Transparency	Conductivity	pH	Total dissolved solids	Dissolved oxygen	UV absorption	Total hardness	Orthophosphates	Total suspended solids	Turbidity	Ambient temperature	Water temperature
True color	1												
Transparency	− 0.481	1											
Conductivity	− 0.097	−0.159	1										
pH	− 0.275	−0.174	−0.074	1									
Total dissolved solids	0.008	−0.147	0.827	−0.288	1								
Dissolved oxygen	− 0.319	−0.048	−0.222	0.751	−0.315	1							
UV absorption	0.373	0.064	0.098	0.044	0.272	0.280	1						
Total hardness	0.052	−0.218	0.441	−0.333	0.396	−0.453	−0.375	1					
Orthophosphates	0.225	0.221	−0.052	−0.241	0.012	−0.309	0.329	−0.043	1				
Total suspended solids	0.849	−0.400	−0.134	−0.417	0.015	−0.294	0.424	0.033	0.443	1			
Turbidity	0.886	−0.464	−0.073	−0.306	0.031	−0.415	0.270	0.094	0.345	0.825	1		
Ambient temperature	0.232	0.047	−0.023	−0.437	−0.011	−0.395	−0.134	0.235	−0.285	0.171	0.110	1	
Water temperature	0.166	−0.286	0.530	−0.295	0.407	−0.390	−0.066	0.246	−0.146	0.174	0.083	0.450	1

Table 5 | Analysis of PCA variances by component (bold entries are variable with highest covariance of each component)

	Component (PC)							
Parameter	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**
Chlorophyll a	0.150	−0.359	0.145	0.426	**0.537**	−0.371	−0.118	0.181
Fecal coliforms	**0.704**	0.024	−0.205	0.399	0.180	0.461	0.041	0.133
Escherichia coli	0.039	−0.181	0.201	0.195	−0.334	**0.619**	0.279	0.344
TOC	0.265	−0.051	0.127	0.365	−0.166	−0.198	−0.128	**−0.724**
BOD	−0.119	−0.474	0.100	**−0.537**	0.376	0.361	−0.212	−0.125
COD	0.234	−0.381	0.203	**−0.588**	0.038	0.225	−0.494	−0.083
Ammoniacal nitrogen	0.415	−0.107	0.135	**−0.535**	−0.417	−0.337	−0.119	0.379
Nitrites	**0.766**	−0.201	−0.081	−0.196	−0.156	−0.331	−0.022	0.172
Nitrates	**0.940**	−0.059	−0.155	−0.017	0.112	0.067	0.072	−0.097
Organic nitrogen	−0.161	**0.924**	−0.154	0.074	0.118	0.024	0.006	−0.105
Total nitrogen	0.555	**0.772**	−0.209	−0.086	0.077	−0.027	0.021	−0.031
Total Kjeldahl nitrogen	−0.057	**0.951**	−0.126	−0.070	0.010	−0.067	−0.027	−0.007
Total phosphorus	**0.939**	−0.088	−0.132	−0.074	−0.136	−0.147	−0.063	0.017
True color	**0.905**	0.118	−0.143	0.200	−0.061	0.203	0.007	−0.093
Transparency	−0.426	−0.523	0.183	−0.146	−0.116	−0.044	**0.533**	−0.229
Conductivity	−0.052	**0.681**	0.450	−0.371	0.223	0.025	0.188	0.089
pH	−0.471	0.322	**−0.738**	0.055	0.083	0.088	−0.033	0.195
Total dissolved solids	0.100	0.462	0.506	**−0.561**	0.269	0.165	0.080	−0.121
Dissolved oxygen	−0.416	0.066	**−0.742**	−0.286	−0.202	−0.103	−0.264	−0.015
UV absorption	0.410	0.040	−0.387	**−0.657**	−0.156	0.158	0.319	−0.212
Total hardness	0.096	0.251	**0.646**	0.105	0.319	−0.038	−0.282	0.128
Orthophosphates	0.443	−0.330	−0.062	−0.137	0.453	−0.334	**0.535**	0.062
Total suspended solids	**0.975**	−0.061	−0.087	−0.059	−0.039	−0.086	−0.067	0.027
Turbidity	**0.876**	0.055	−0.099	0.171	0.226	0.231	−0.045	−0.050
Ambient temperature	0.218	0.029	0.513	0.257	**−0.572**	0.135	−0.080	−0.033
Water temperature	0.238	**0.637**	0.474	0.042	−0.333	−0.195	0.116	0.034

adsorption-desorption processes, depending on amount, unit, type and contact time between particulate matter and water. Particulate matter originates from organic matter, nutrients and toxic organic and inorganic pollutants, with natural or anthropogenic sources (Chapman 1996).

The PCA parameters found with the highest variance belong to the particulate matter group included in PC1, PC2, PC4 and PC6 (Table 5): total suspended solids (0.975), true color (0.905), turbidity (0.876), fecal coliforms (0.704), *Escherichia coli* (0.619), total dissolved solids (0.561) and transparency (−0.523). According to Pearson correlation (Table 4), total suspended solids showed a positive correlation with turbidity (0.825) and true color (0.849). For its part, fecal coliforms also showed strong correlation with turbidity (0.842) and true color (0.814), as well as a negative correlation with transparency (−0.446). Turbidity, on the other hand, showed a significant positive correlation (0.886) with true color while true color showed a moderate negative correlation (−0.481) with transparency. Transparency decreases as true color, suspended sediments, bacteria, phytoplankton, algae abundance and other organisms increase, these parameters are directly related to increased turbidity in water (EPA 2006). In this way, parameters such as true color, turbidity and transparency were eliminated because these are related with total suspended solids and fecal coliforms. Although true color has a high variance of 0.905 in PC1 (Table 5), this parameter was not considered because it is a dependent variable with high positive correlation with total suspended solids, fecal coliforms and turbidity, in addition to having a negative mean correlation with transparency (−0.481) (Table 4). Total dissolved solids (0.561) and *Escherichia coli* (0.619) were eliminated from PC4 and PC6, respectively, because fecal coliforms and total suspended solids had greater variance in PC1 of the PCA result, which means a greater representativeness in the parameters chosen with related characteristics (Table 5). On the other hand, according to the descriptive statistical results (Table 2), total suspended solids have a higher percentage of variation (standard error/mean = 35.21%) compared with total dissolved solids (2.93%) (Table 3).

Fecal coliforms is a group related to bacterial contamination elements with the highest significant impact on human health. The presence of fecal coliforms in bodies of water indicates fecal contamination and possibly the presence of other organisms including *Escherichia coli, Enterobacter, Klebsiella*, and *Citrobacter* (Tufail *et al.* 2008). Therefore, fecal coliforms serves as an indicator of livestock activity upstream ALMD. The other significant parameter identified was the total suspended solids, which represents waste coming to the system from physical and chemical processes involved in rock weathering, soil erosion and sediment transport in the reservoir. The Environmental Protection Agency considers total suspended solids and fecal coliforms (pathogens) as basic pollutants defining water quality in natural reservoirs (EPA 2006). In this sense, total suspended solids and fecal coliforms were parameters selected to relate to particulate matter group.

Nitrogen compounds

Nitrogen is essential for living organisms in all their forms. However the excess or lack of nitrogen, in relation to phosphorus, define overproduction or nutrient deficiency which can lead to significant contamination levels or eutrophication related to anoxic conditions in the system (Nikitin *et al.* 2015). Nitrogen-based compounds were classified in PC1 (nitrates 0.940 and nitrites 0.766), PC2 (total Kjeldahl nitrogen 0.951, organic nitrogen 0.924 and total nitrogen 0.772) and PC4 (ammoniacal nitrogen −0.535) (Table 5). Nitrates were found with a higher variance coefficient in PC1, (0.940) compared with nitrites (0.766) (Table 5) in agreement with descriptive statistical analysis where nitrates represent a higher variance percentage (35.25%) than nitrites (27.0%) (Table 2). Nitrites showed moderate correlation with nitrates (0.666) and ammoniacal nitrogen (0.638) being the only nitrogen-based compound in PC4 (Table 4). For this study, it is worth of note that nitrates behave as the most oxidized chemical species in the nitrogen cycle, often occurring in shallow water bodies, in comparison to nitrites because bacterial transformation of nitrites is fast (Xia *et al.* 2015). Due to the importance of nitrates in reservoir eutrophication, and algae production as well as its significant statistical variation, nitrites were eliminated. Ammoniacal nitrogen (variance −0.535), the most reduced form of nitrogen, was selected as the only nitrogen compound in PC4 (Table 5). The ammoniacal nitrogen showed no significant correlation with the same type of compounds as in the Pearson correlation (Table 4). PC2 shows a strong influence of nitrogen-based compounds with the highest variance of total Kjeldahl nitrogen (0.51) followed by organic nitrogen (0.924) (Table 5). Organic nitrogen was selected even though total Kjeldahl nitrogen showed a higher variance coefficient in the PCA. Organic nitrogen had a significant positive correlation (0.966) with total Kjeldahl nitrogen and total nitrogen (0.700) (Table 4). Based on the descriptive statistical analysis, organic nitrogen showed a higher variation percentage (8.1%) than total nitrogen (6.3%) and total Kjeldahl nitrogen (6.1%) (Table 2). Because inorganic nitrogen occurs in the environment as nitrate, nitrite, ammonium ion, and molecular nitrogen, total Kjeldahl nitrogen and total nitrogen were not considered because they are represented by nitrates and ammoniacal nitrogen and have significant correlation with organic nitrogen. Organic nitrogen represents proteins, amines and amino acids present in the system and their availability for biochemical transformation in the food chain by phytoplankton and/or bacteria, and considerable increases in this parameter may indicate contamination (Zieliński *et al.* 2013).

Physico-chemical

There are general variables that influence the vital processes of organisms according to the behavior and physical and chemical transformations that occur in the ecosystem. The parameters identified as general variables according to the PCA were water and air temperature, electric conductivity, hydrogen potential (pH), dissolved oxygen and total hardness. Dissolved oxygen is one of the most important parameters in this category and a key factor for aquatic life development and the reservoirs' chemical characteristics (Aiping *et al.* 2015). Dissolved oxygen was selected because of its variance coefficient value (−0.416) and because it was the only parameter of this category classified in PC1 (Table 5). Dissolved oxygen varies with temperature, salinity, turbulence, photosynthetic activity of algae and plants as well as atmospheric pressure. The solubility of oxygen in water is reduced when temperature and salinity increase. Hardness in water depends mainly on the presence of dissolved calcium and magnesium salts (Chapman 1996). Considering these relationships and according to the Pearson correlation (Table 4), dissolved oxygen is a variable depending on air (−0.395) and water (−0.309) temperature and total hardness (−0.453) and thus, these last three parameters were not considered. Average water (29.6 °C) and air temperature (33.6 °C) values showed a maintained low percentage of variation in the dataset (1.49% and 1.27%, respectively) (Table 2), meaning there is no significant variation in temperature values. Electric conductivity was not considered as a parameter because it is associated with the concentration of filterable matter and the presence of some ions, is an approximate indicator of mineral content and closely related to total dissolved solids (parameter also not considered), with significant correlation (0.827) (Table 4).

pH is considered a highly significant parameter in the water column because it is related to the degree of acidity or alkalinity in the aquatic ecosystem. When pH is altered, the biological processes are disturbed or inhibited by changes in the aerobic or anaerobic zone distribution, or in the redox processes occurring at the thermocline. In other reports, dissolved oxygen and pH are considered essential for understanding the dynamics of other physicochemical contaminants and primary productivity and their effect on water quality in a reservoir (Rajendran & Mansiya 2015). In this work, pH was considered the only parameter with the highest variance (−0.738) in PC3 (Table 5), and because total hardness was not considered, pH was selected.

A different classification for physicochemical group corresponds to compounds derived from organic matter: chemical oxygen demand (COD), biochemical oxygen demand (BOD) and total organic carbon (TOC). Several natural reservoirs contain organic matter that can be measured as TOC. However, for water quality comparative purposes, BOD and COD have been used as indicators to measure organic matter concentration in a reservoir. COD includes biodegradable and non-biodegradable organic and inorganic material, unlike BOD which considers only biodegradable matter (Lee *et al.* 2016), therefore: COD > BOD > TOC. Although TOC was found with greater variance (−0.724) than COD (−0.588) and BOD (−0.537), COD was classified as a component with the greatest data representativeness (Table 5). Regarding the correlation analysis, BOD presents a significant correlation (0.706) with COD (Figure 4). In this way, COD was selected to consider biodegradable and non-biodegradable compounds because agricultural activities are carried out within the study area, including application of pesticides, as well as mining activity upstream that may generate pollution of the reservoir.

Phosphorus compounds

Phosphorus is usually the limiting nutrient influencing algae growth and controlling primary productivity in a reservoir, usually dissolved as orthophosphates and polyphosphates. Their main sources are rock weathering and organic matter decomposition (Zaragüeta & Acebes 2017). The resulting parameters from this category are: total phosphorus, representativeness coefficient 0.939 at PC1 (one of the highest), and orthophosphates (0.535), only in the PC7 (Table 5). High phosphorus concentration not only triggers primary productivity, but also defines the degree of contamination of a reservoir. In this regard, two forms of phosphorus were included for WQI_{ALMD}, to verify the trophic state of the system: total phosphorus and orthophosphates in order to include both inorganic and organic phosphorus, because in tropical water bodies, phosphorus is the nutrient that limits the system and excessive levels can lead to serious problems in the reservoir (Li & Nan 2017).

Primary productivity

Zooplankton and phytoplankton play a significant role in the reservoir's food chain through the photosynthesis process generated from the production of chlorophyll in plants and the absorption of solar UV energy (Chen *et al.* 2017). Contamination by primary productivity originates from organic and inorganic matter and biomass produced in microorganisms by carbon, nitrogen and phosphorus transformation. The excessive presence of nutrients, mainly phosphorus and nitrogen, are indicators of the trophic state of a reservoir. The concerning parameters in this category are: UV absorption (variance −0.657) classified in PC4, and chlorophyll a (variance 0.537) classified in PC5 (Table 5). Chlorophyll a was the last selected parameter with high variance coefficient (0.537) in PC5 (Table 5). Chlorophyll a showed average correlation with UV absorption (−0.481) (Table 4) and higher variation (17.0%) than UV absorption (8.6%) (Table 2) probably because UV absorption is an indirect measure of chlorophyll a. Chlorophyll a indicates the evolution of the main components of the trophic chain in a water body, indispensable for life and maintenance of the reservoir. A high chlorophyll a concentration, however, indicates eutrophication processes in the reservoir (Nikitin *et al.* 2015). In this sense, chlorophyll a was included to address an assessment from the point of view of the trophic status.

Calculation of WQI_{ALMD}

The 11 parameters described in this section were selected to be part of the WQI_{ALMD} because they are the most statistically significant (PCA and Pearson correlation) for evaluating water quality in the reservoir. Once these parameters were identified, Equation (1) was used to estimate WQI_{ALMD}. For calculating *Sli* in Equation (1), every parameter was individually analyzed according to their probability functions (Figure 3). For determination of intrinsic weights, representative PCA values were considered. First, intrinsic weight was assigned to the selected variables

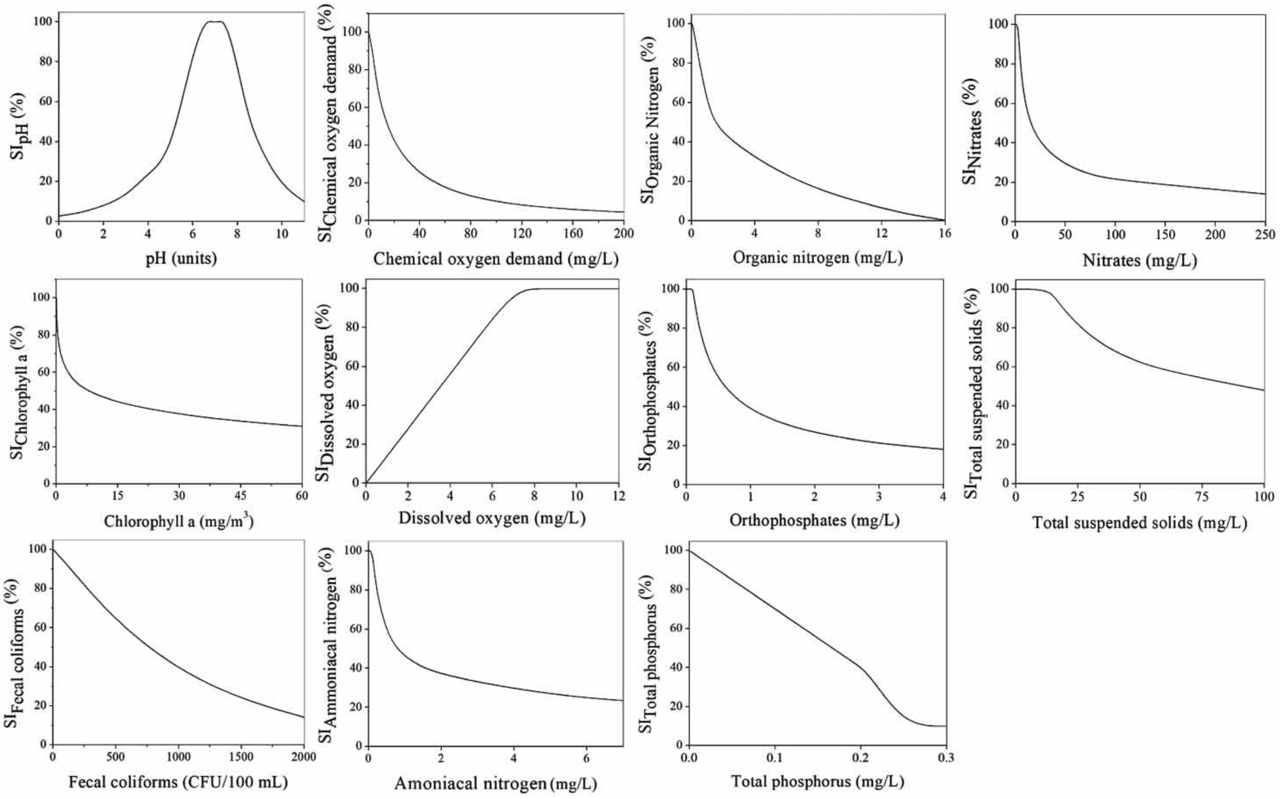

Figure 3 | Quality functions for the calculation of *SIi*.

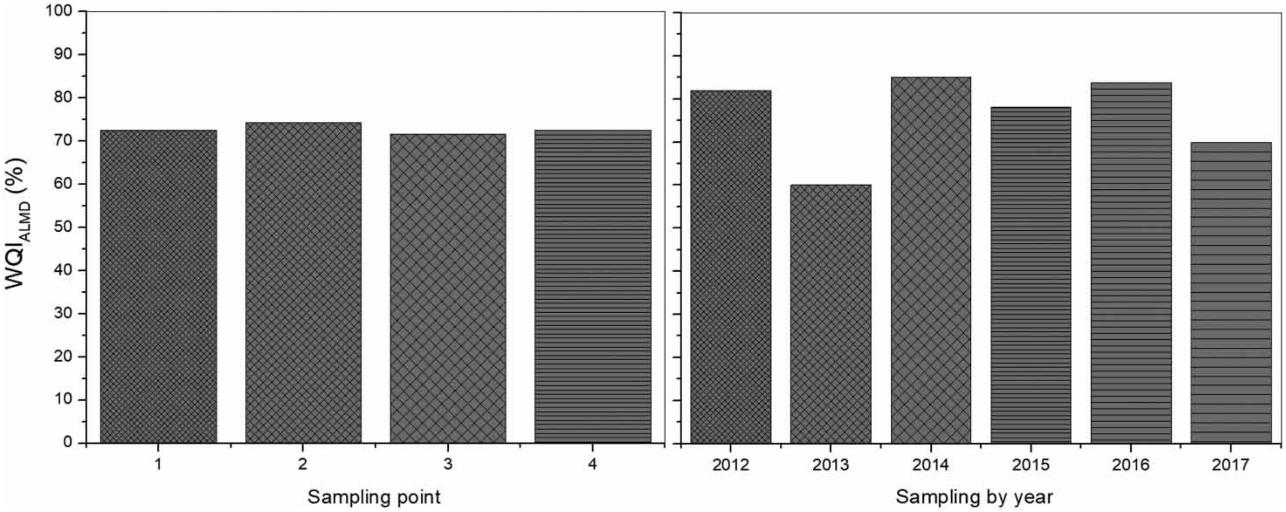

Figure 4 | Spatial and temporal variation of WQI_{ALMD} (2012–2017).

representing each component itself, being fixed parameters of *Wi*. For example, Table 6 shows that the orthophosphates intrinsic weight is 0.05, chlorophyll a 0.07 and organic nitrogen 0.17.

For the remaining parameters, an algorithm was established to generate random *Wi* values, building restrictions into the system based on the PCA results. These restrictions were: dissolved oxygen + pH = 0.12; COD + ammoniacal nitrogen = 0.1; organic nitrogen = 0.17; chlorophyll a = 0.07; orthophosphates = 0.05. The maximum value of the sum of all the weights must be 1.0, and the minimum value assigned to the parameters should be 0.01 because lower values are not significant for the index (Hajkowicz 2006). With all of these restrictions, the Simplex Method generated 65,532 possible solutions. In this sense, the *Wi* corresponding to the average (WQI_{ALMD} simulated) represents the highest probability of data adjustment that responds to the mathematical model for WQI_{ALMD} avoiding *Wi* subjective selection.

Table 6 | Calculation of the WQI_{ALMD}

Parameter	*Wi*	*Sli*	General WQI_{VALUE} *(Wi*Sli)*
Fecal coliform	0.26	49.64	12.91
Total suspended solids	0.03	71.51	2.15
pH	0.06	73.76	4.43
Dissolved oxygen	0.06	94.65	5.68
COD	0.01	37.64	0.38
Organic nitrogen	0.17	73.06	12.42
Ammoniacal nitrogen	0.09	100	9.00
Nitrates	0.14	100	14.00
Total phosphorus	0.06	67.95	4.08
Orthophosphates	0.05	100	5.00
Chlorophyll a	0.07	49.82	3.49
WQI_{ALMD}			73.52

Optimal weights selected for the model obtained were: fecal coliforms 0.26; total suspended solids 0.03; pH 0.06; dissolved oxygen 0.06; COD 0.01; nitrates 0.14; organic nitrogen 0.17; ammoniacal nitrogen 0.09; total phosphorus 0.06 orthophosphates 0.05; and chlorophyll a 0.07. Once obtained the *WI* and *Sli* for all the parameters, the WQI_{VALUE} values were calculated to determine a general WQI_{ALMD} with all selected database parameters (Table 6).

According NSF (Brown *et al.* 1970) classification for water quality indexes with multiplicative sum features, the classification scale used is: excellent (91–100), good (71–90), medium (51–70), poor (26–50) and very bad (0–25).

The general WQI_{ALMD} value estimated was 73.52 and the water quality classification for ALMD is ‘good’. The classification obtained responds to the parameters relating to the eutrophication of the system in the ALMD.

In the results of each WQI_{VALUE}, the most critical variables to define the quality of water at ALMD are: COD

(0.37), total suspended solids (2.03), chlorophyll a (3.29) and total phosphorus (3.92) because they have lower values. These critical variables respond to phosphorus compounds, of primary productivity, physico-chemical and suspended matter. On the other hand, the variables with the greatest value are nitrates (14.00), organic nitrogen (12.36), fecal coliforms (11.66) and ammoniacal nitrogen (9.00), responding mostly to compounds of nitrogen. The result of the index from a general perspective allowed for classification of the water as good quality (Rajendran & Mansiya 2015). In order to obtain a temporal and spatial evaluation of the reservoir, the proposed index was calculated for each year and sampling point (Figure 4).

The results obtained in spatial variation of ALMD indicate that the dam has maintained a water quality classified as 'good' for the four sampling sites in the period 2012–2017. Site 2 ($WQI_{ALMD} = 74.25$) had the highest index, while the three remaining sites had very similar indices (site 1: $WQI_{ALMD} = 72.46$, site 3: $WQI_{ALMD} = 71.63$ and site 4: $WQI_{ALMD} = 72.45$). The spatial result showed that there are no significant variations on aspects related to the water quality of the reservoir in relation to the sampling sites 1, 3 and 4 (Figure 4(a)). Sampling site 2 has a higher index due to the decrease in chlorophyll a (7.62 mg/m^3) and fecal coliform concentrations (602 CFU/100 mL) compared with site 3 which has a lower index (chlorophyll a = 11.36 mg/m^3; fecal coliforms = 851 CFU/100 mL). Changes in spatial variations in ALMD are mainly attributed to variables related to primary productivity and the increased bacteria.

Annual indices indicate that ALMD had 'good' water quality in 2012 ($WQI_{ALMD} = 81.94$), 2014 ($WQI_{ALMD} = 85.00$), 2015 ($WQI_{ALMD} = 78.06$) and 2016 ($WQI_{ALMD} = 83.79$). However, the reservoir shows significant variations in 2013 and 2017. The most representative parameters of the ALMD were affected in 2013 and 2017, as the indices decreased considerably due to the increase in pollutant concentrations. Averages of the concentrations in 2013 (chlorophyll a = 18.11 mg/m^3, fecal coliforms = 2,127 CFU/100 mL, total phosphorus = 0.24 mg/L, dissolved oxygen = 5.07 mg/L and total suspended solids = 146 mg/L) were higher than in 2014 (chlorophyll a = 4.03 mg/m^3, fecal coliforms = 221 CFU/100 mL, total phosphorus = 0.06 mg/L, dissolved oxygen = 7.28 mg/L and total suspended solids = 9.53 mg/L). The same situation occurred from 2016 to 2017 with a lower proportion. However, the increase in pollutant concentration from 2012 to 2013 and 2016 to 2017 originated due to variables attributed to eutrophication in the system (Doan *et al.* 2015). The decrease of 21.96% in 2012 ($WQI_{ALMD} = 81.94$) to 2013 ($WQI_{ALMD} = 59.98$) caused a change in the classification scale used (Rajendran & Mansiya 2015), from 'good' to ' medium' quality. As in 2013, the reservoir presents 'medium' water quality in 2017 due to the reduction of the index by 13.91% of 2016 ($WQI_{ALMD} = 83.79$) to 2017 ($WQI_{ALMD} = 69.88$). The WQI_{ALMD} value of the best quality ($WQI_{ALMD} = 85.00$) was found in 2014, while the lowest value ($WQI_{ALMD} = 59.98$) was obtained in 2013 attributed to the increase in pollutant concentrations (Figure 4(b)).

Sensitivity analysis

For this analysis variations of ±30% and ±50% were performed for each parameter that forms the general WQI_{ALMD}, generating values for each Sli and Wi as input variables. The variation of ±30% is the usual confidence interval for this type of calculation. The interval ±50% was selected to ensure the sensitivity of the model. The output variables were variations of the results (±30% and ±50%) obtained for each WQI_{ALMD}. The simulated WQI_{ALMD} values fluctuated between 65 and 79 with a standard deviation of 1.5335 and a variation of 3% in model. Once the inputs and outputs were identified, the SC of each parameter was calculated for each percentage of variation, the results are shown in Figure 5.

The proposed model presents greater fluctuation sensitivity to the percentage change in positive values (+30% and +50%). The most sensitive variables in the model were: orthophosphates (1.00 = +50%, 0.88 = +30%), dissolved oxygen (0.99 = +50%, 0.87 = +30%), ammoniacal nitrogen (0.97 = +50%, 0.86 = +30%) and nitrate (0.94 = +50%, 0.85 = +30%). These variables respond to the group of nitrogen and phosphorus compounds, as well as the physico-chemical parameters relating to the depletion of oxygen. With regard to the negative fluctuations (−30% and −50%), the most sensitive variables were the same as those of the positive fluctuations, but had a lower sensitivity.

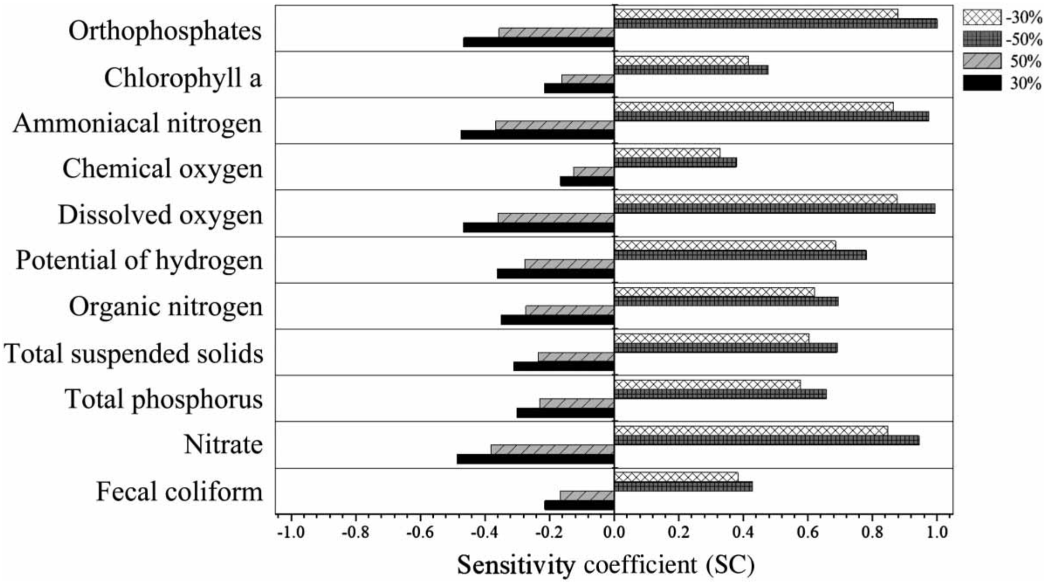

Figure 5 | Sensitivity analysis of model parameters.

CONCLUSIONS

The WQI_{ALMD} obtained represents the quality of the water of an important tropical reservoir in the region of Sinaloa, the state with the greatest agricultural production in Mexico. This index represents the state of the water quality in the period of the last 5 years. The analysis of main components and the Pearson correlation matrix of the variables were determined with greater representation in the water body corresponding to physical and chemical characteristics of suspended matter, nitrogen and phosphorus compounds, as well as factors that intervene in the primary productivity of the reservoir. In spite of insufficient scientific information for this reservoir, the statistical tools used assisted with algorithms in programming, and showed its potential for the construction of a WQI particularly in tropical environments. It is recommended to carry out a specific study on heavy metals and pesticides in the system, since the result of the proposed index shows that the main contaminants present in the dam are related with chemical compounds and phosphorus, which can be the by-product of mining originating upstream of the dam and excessive use of pesticides in agricultural crops, respectively. Given the fact that there are other dams in the region where only the parameters of water quality are registered, the methodology proposed in the present paper would be of great relevance in obtaining the respective WQI of each one of the dams and would inform the agencies involved in the decision-making process, thereby improving the conservation of water resources of Sinaloa and thus, helping to guarantee food supply at a national and international level.

ACKNOWLEDGEMENTS

The authors recognize the contribution of CONAGUA for the availability of the database to support the realization of this work. The authors also acknowledge the National Council of Science and Technology in Mexico (CONACYT) for supporting this research work.

REFERENCES

Aiping, T., Jinbao, W., Rong, W., Shuang, L. & Hongyan, S. 2015 Importance of pH, dissolved oxygen and light to phosphorus release from ditch sediments. *Nature Environment and Pollution Technology* **14**, 475–484.

An, S., Xie, X. & Ma, Y. 2015 Evaluation of water quality using principal component analysis. *Nature Environment and Pollution Technology* **14**, 855–858.

Bautista, J. & Ruiz, J. 2011 *Calidad de Agua Para el Cultivo de Tilapia en Tanques de Geomembrana (Water Quality for the Cultivation of Tilapia in Geomembrane Tanks)*. Direction of Strengthening Research, Autonomous University of Nayarit, Mexico. http://fuente.uan.edu.mx/publicaciones/03-08/2.pdf (Accessed 20 July 2017).

Brown, R. M., McClelland, N. I., Deininger, R. A. & Tozer, R. G. 1970 A water quality index – Do we care? *Water & Sewage Works* **117**, 339–343.

Chapman, D. 1996 *Water Quality Assessments. A Guide to the Use of Biota, Sediments and Water in Environmental Monitoring*. Chapman & Hall, London.

Chen, H., Zhou, W., Chen, W., Xie, W., Jiang, L., Liang, Q. & Wang, Q. 2017 Simplified, rapid, and inexpensive estimation of water primary productivity based on chlorophyll fluorescence parameter F_o. *Journal of Plant Physiology* **211**, 128–135.

Cude, C. 2001 Oregon water quality index: a tool for evaluating water quality management effectiveness. *Journal of the American Water Resources Association* **37**, 125–137.

Doan, P. T. K., Némery, J., Schmid, M. & Gratiot, N. 2015 Eutrophication of turbid tropical reservoirs: scenarios of evolution of the reservoir of Cointzio, Mexico. *Ecological Informatics* **29**, 192–205.

Environmental Protection Agency (EPA) 2006 Chapter 15, Turbidity and Total Solids, Voluntary Estuary Monitoring Manual, March 2006. https://www.epa.gov/sites/production/files/201509/documents/2009_03_13_estuaries_monitor_chap15.pdf (Accessed 20 July 2017).

Ferré, L. 1995 Selection of components in principal component analysis: a comparison of methods. *Computational Statistics and Data Analysis* **19**, 669–682.

Hajkowicz, S. 2006 Multi-attributed environmental index construction. *Ecological Economics* **57**, 122–139.

Howladar, M., Numanbakth, M. & Faruque, M. 2017 An application of Water Quality Index (WQI) and multivariate statistics to evaluate the water quality around Maddhapara Granite Mining Industrial Area, Dinajpur, Bangladesh. *Environmental System Research* **6**, 1–18.

Lee, J., Lee, S., Yu, S. & Rhew, D. 2016 Relationships between water quality parameters in rivers and lakes: BOD_5, COD, NBOPs, and TOC. *Environmental Monitoring and Assessment* **188** (4).

Li, X. & Nan, R. 2017 A bibliometric analysis of eutrophication literatures: an expanding and shifting focus. *Environmental Science and Pollution Research* **24**, 17103–17115.

Moscuzza, C., Volpedo, A., Ojeda, C. & Fernández, A. 2007 Water quality index as a tool for river assessment in agricultural areas in the Pampean plains of Argentina. *Journal of Urban and Environmental Engineering* **1**, 18–25.

National Water Commission (CONAGUA) 2015 System for monitoring dams. http://conagua.presas.mx/ (Accessed 19 March 2017).

Nikitin, O., Stepanova, N. & Latypova, V. 2015 Human health risk assessment related to blue-green algae mass development in the Kuibyshev reservoir. *Water Science and Technology* **15** (4), 693–700.

Palácio, S. M., Espinoza-Quiñones, F. R., de Pauli, A. R., Piana, P. A., Queiroz, C. B., Fabris, S. C. & Veit, M. T. 2016 Assessment of anthropogenic impacts on the water quality of Marreco river, Brazil, based on principal component analysis and toxicological assays. *Water, Air and Soil Pollution* **227**, 1–11.

Rajendran, A. & Mansiya, C. 2015 Physico-chemical analysis of ground water samples of coastal areas of south Chennai in the post-tsunami scenario. *Ecotoxicology and Environmental Safety* **121**, 218–222.

Rangel-Peraza, J., De Anda, J., Gonzales, F. & Rode, M. 2016 Sensitivity and uncertainty analysis on water quality modelling of Aguamilpa reservoir. *Journal of Limnology* **75**, 81–92.

Şener, E., Şener, Ş. & Davraz, A. 2017 Evaluation of water quality using water quality index (WQI) method and GIS in Aksu river (SW-Turkey). *Science of the Total Environment* **584–585**, 131–144.

Tirkey, P., Bhattacharya, T. & Anushree, C. 2015 Water quality indices – important tools for water quality assessment: a review. *International Journal of Advances in Chemistry* **1**, 15–29.

Tufail, M., Ormsbee, L. & Teegavarapu, R. 2008 Artificial intelligence-based inductive models for prediction and classification of fecal coliform in surface waters. *Journal of Environmental Engineering* **134**, 789–799.

Varol, S. & Davraz, A. 2015 Evaluation of the groundwater quality with WQI (water quality index) and multivariate analysis: a case study of the Tefenni plain (Burdur/Turkey). *Environmental Earth Sciences* **73**, 1725–1744.

Xia, L. Z., Liu, G. H., Wu, Y. H., Ma, L. & Li, Y. D. 2015 Protection methods to reduce nitrogen and phosphorus losses from sloping citrus land in the Three Gorges area of China. *Pedosphere* **25**, 478–488.

Zaragüeta, M. & Acebes, P. 2017 Controlling eutrophication in a Mediterranean shallow reservoir by phosphorus loading reduction: the need for an integrated management approach. *Environmental Management* **59**, 635–651.

Zieliński, R., Dunalska, J., Grochowska, J., Bigaj, I. & Szymański, D. 2013 Variation of nitrogen forms in lakes with different intensity of anthropogenic pressure. *Limnological Review* **13**, 181–188.

First received 22 February 2018; accepted in revised form 13 September 2018. Available online 10 October 2018