

## Sediment Characterization of Unlined Water-canal in Chókwè Irrigation Scheme, Mozambique Using Multivariate Statistical Analysis

Lateiro Salvador de Sousa<sup>1\*</sup>, James M Raude<sup>2</sup>, Raphael M Wambua<sup>3</sup> and Benedict M Mutua<sup>4</sup>

<sup>1</sup>Division of Agriculture, Instituto Superior Politécnico de Gaza, Chókwè-Gaza Province, Mozambique.

<sup>2</sup>Soil, Water & Environmental Engineering Department, Jomo Kenyatta University of Agriculture & Technology, Nairobi, Kenya.

<sup>3</sup>Department of Agricultural and Biosystems Engineering, South Eastern Kenya University, Kitui County-Kenya.

<sup>4</sup>Office of Vice-Chancellor, The Technical University of Kenya, Kenya.

### \*Correspondence:

Lateiro Salvador de Sousa, Division of Agriculture, Instituto Superior Politécnico de Gaza, Chókwè-Gaza Province, Mozambique, ORCID: 0000-0001-9375-3826.

E-mail: lateiro.sousa@ispg.ac.mz or lateirodesousa@gmail.com

Received: 09 Mar 2025; Accepted: 22 Apr 2025; Published: 30 Apr 2025

**Citation:** de Sousa LS, Raude JM, Wambua RM, et al. Sediment Characterization of Unlined Water-canal in Chókwè Irrigation Scheme, Mozambique Using Multivariate Statistical Analysis. Int J Agriculture Technology. 2025; 5(2): 1-12.

### ABSTRACT

Chókwè Irrigation Scheme (CIS) has a crucial role of enhancing agricultural production and food security in Gaza Province in Mozambique. However, sediment deposition in the canal system in recent past has adversely affected water conveyance and distribution efficiency. To formulate best management practices for the sediment, there is need to understand its spatial characteristics and distribution. The objective of this research was to characterize sediment distribution in CIS using multivariate statistical analysis (MSA) technique. Variables were divided into two groups; canal channel factors (CCF) and water inflow factors (WIF), with respectively, twelve (12) and eleven (11) physicochemical parameters analyzed for nine (9) sampling stations. The sampling procedures were observed for six (6) months and two (2) seasons, respectively, during the months of June and August in 2018, for dry season (DS) and, of January and March in 2019, for the wet season (WS). Results indicated that F1 for CCF in the DS, explained 33.24% in variance and provided strong correlation towards water depth, canal depth, critical shear stress, plasticity index, electrical conductivity, exchangeable sodium percentage (ESP), concentrations of sodium, and combination of calcium and magnesium; whilst weak correlation was found for sediment settling velocity. During WS, F1 explained 48.09% in variance and provided strong correlation leaning to water and canal depths, critical shear stress, plasticity index, electrical conductivity, ESP, sodium adsorption ratio (SAR), concentrations of sodium, potassium, and combined calcium and magnesium. Conversely, for the WIF, F1 explained 30.59% for DS and 44.85 for WS. Sampling stations distinctly comprised the main clusters, both for CCF and WIF factors, both for DS and WS. The study revealed that canal channel factors and inflow factors appear to be behind sedimentation occurrence at CIS, with geometric, water inflow-flux and physicochemical parameters explaining the causes. These results may be useful in irrigation system management for improved agricultural productivity and enhanced food security.

### Keywords

Sediment Characterization, Multivariate Statistical Analysis, Principal Component Analysis, Cluster Analysis, Unlined Water-Canal, Canal channel factors and Water inflow factors, Chókwè Irrigation Scheme.

### Introduction

Water conveyance in an irrigation scheme, such as of the Chókwè Irrigation Scheme (CIS), aims at water application in an efficient manner that is, the timely application of water and its proper regulation to minimize the wastage of water, soil, and the nutrients in the crops [1]. The irrigation scheme is set to use water for

irrigating cultivated plants with profitability and sustainability to augment crop yield [2] and improve human being livelihoods at the same time reducing food insecurity [3]. In addition, it can also be used to control water usage and supply to farmed lands with crops in accordance with their needs at distinct growth phases. This fact becomes critical when soil erosion control, soil salinity minimization, and reduction of groundwater pollution is taken into consideration [4-6]. Steady and stable water delivery can be achieved by properly designed and maintained irrigation canals [7,8], where various hydraulic parameters are taken into consideration, including water and sediment flow factors [9-13]. Good understanding of the interaction between water and sediment flow factors in a given irrigation canal can be analyzed with resource to multivariate statistical methods [14-18], that includes Principal Component Analysis (PCA) [16]. Different research work has taken place worldwide seeking to make contribution into the field of knowledge for water quality improvement in rivers and other water body systems, including canal irrigation system using multivariate statistical approach.

Multivariate statistical analysis has widely been used for clarification and reduction of multifaceted dataset of water quality characteristics [16,19]. These techniques can be used to generate easily interpretable results, helping in that manner to assess deterioration of water quality [18] and by extension, sedimentation effects of a given water body [20,21]. A study by Fan et al. [22] on the agricultural lands polluted by certain amount of heavy metal and the consequent risks imposed to human health in the southern margin of Tarim Basin, Xinjiang, in China, used multivariate statistical analysis (MSA). Another study by Hui et al. [23] applied index of water quality and MSA for the assessment of the hydro-geochemistry analysis of shallow groundwater in Hailun, northeast China. Kabir et al. [24] used PCA and cluster analysis (CA) to study spatial and seasonal variation of the water quality and its suitability at the Shitalakhya river, which its surroundings in the urban river in Bangladesh has been characterized to be economically vibrant but ecologically challenging, due to different factors associated to sources of water pollution in the region.

A study by Li et al. [25] on the assessing and sourcing the identification of the toxic metals in an abandoned synthetically produced diamond plant from Anhui Province, China, successfully used MSA for the mine generated acid wastewater. Another study conducted by Liu et al. [26] on a multivariate statistical technique used a comprehensive analysis, computing analysis of variance, correlation analysis, principal component analysis and factor analysis, and multiple linear regression model, to estimate spatial-temporal changes in 13 parameters for water quality of Ganwol reservoir. In another study on sediment dynamics multivariate analysis was used on the Tarakan sub-basin, North Kalimantan, Indonesia. MSA proved to be reliable when used to identify sediment particles having similarities in their features and showed to be a good representation to distinguish sedimentary faces and environmental structures of deposition. Furthermore, in another study performed by Prajapati et al. [27], MSA demonstrated to be useful for hierarchic clustering analysis as well as for factor

analysis of the geochemical parameters of groundwater samples, from which hidden sources information were made known. As shown in these previous studies, MSA appears to be a useful tool in extracting the most important parameters for water quality analysis, as this is true for PCA [28-30]. PCA can be used to identify factors responsible for water quality variations at different and broad scale [31]. Therefore, recent studies revealed successfully, strong relation of MSA to specific contamination sources [21,32-34], which can also be applied for sedimentation study.

Cluster Analysis (CA) is a method in multivariate statistics which is used to group sampling locations based on their level of likenesses in each class or divergences within these same classes [35-37]. The utility of cluster analysis in the assessment of the spatial-temporal changes in the water resources quality management, including sedimentation, and have been thoroughly described [37,38]. Some examples of studies conducted on the spatial-temporal evaluation on the water and sediment flow in Mozambique where MSA was used remain scarce. The CIS has experienced water flow challenges to cope with its goal, of conveying enough water to the most downstream sites of the schemes for a successful crop production. Sedimentation has been reported as the main causes of bathymetric changes in the main canal bed [39]. The decrease in water conveyance capacity in some points of the CIS main canal has been more evident due to the improper management practices which has caused changes in canal channel factors (CCF) and water inflow factors (WIF) relevant to water distribution. CCF and WIF are related to the physio-chemical parameters of both, canal, and water quality. The decrease in water conveyance, and lack of proper sediment management has led to problems of water distribution and low agricultural productivity. To develop best sediment management practices, the sediment characteristics along the canal. However, there is scanty information on spatial sediment characteristics in CIS canal systems. This article aimed at characterizing sediment distribution in CIS using multivariate statistical analysis (MSA) technique.

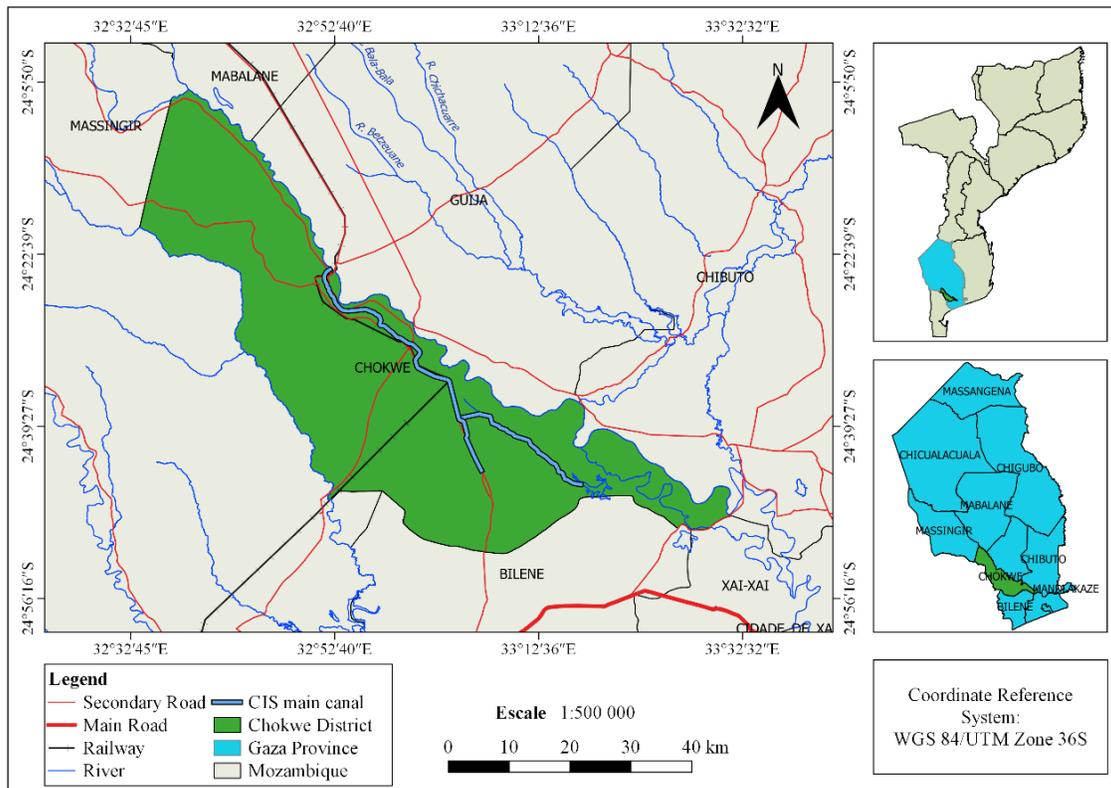
## Materials and Methods

### Study Area

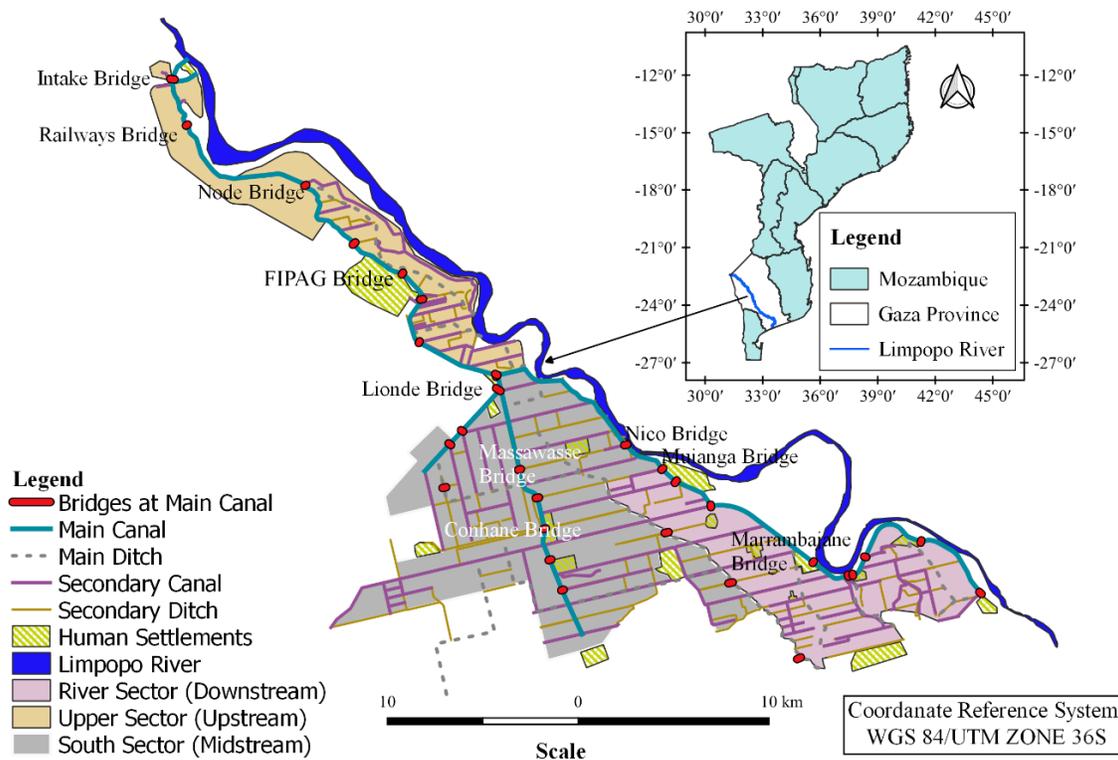
The study was realized in the Chókwe Irrigation Scheme (CIS), situated at the Lower Limpopo River Basin (LLRB), in the Chókwe District, within Gaza Province, in Mozambique. LLRB covers approximately 84 981 km<sup>2</sup>. The study site area is located within latitudes 24°04'32'' and 25°01'35'' South and longitudes 32°40'11'' and 33°37'14'' East. The area is vastly dry where average rainfall is approximately 530 mm/year. The rainfall occurs mostly from October to March in a normal hydrological year, and between December and February during very dry years. Figure 1 presents site map location for the CIS in Mozambique.

### Data Collection

The CIS is situated in Mozambique, and sources water in an approximate average of 45 m<sup>3</sup>/s from the mainstream of the Limpopo River. Deviated water at the main headworks flows into the main canal in the scheme to benefit nearly 12,000 producers. Total CIS area is of 33,000 hectares approximately, mostly for



**Figure 1: Geographical situation for Chokwe Irrigation Scheme in Gaza Province, Mozambique.**



**Figure 2: Chokwe Irrigation Scheme network with sampling stations presented.**

**Table 1:** Sampling Stations and Cumulative Distances from the Intake.

Sampling station code	Station name	Cumulative Distance from Intake (km)	Geographical location		
			Latitude	Longitude	Elevation* (m)
P0G	Intake	0.000	S24°24'12.10"	E032°52'05.90"	42
P1G	Railways bridge	1.527	S24°24'53.21"	E032°51'37.89"	41
P1O	Node	14.102	S24°28'23.29"	E032°56'37.34"	36
P1D	Node bridge	14.154	S24°28'25.30"	E032°56'37.80"	36
P3D	Chókwè bridge (FIPAG)	23.359	S24°31'26.30"	E033°00'15.40"	33
P2O	River canal	33.889	S24°34'55.70"	E033°03'45.62"	31
P6D	Lionde bridge	33.960	S24°34'57.30"	E033°03'47.50"	31
P7D	Massawasse bridge	40.031	S24°38'08.40"	E033°04'39.10"	26
P9D	Conhane bridge	44.312	S24°40'10.10"	E033°05'35.60"	25
P2O	River canal	33.889	S24°34'55.70"	E033°03'45.62"	31
P2R	Nico bridge	44.339	S24°37'28.30"	E033°08'60.90"	27
P3R	Muianga bridge	47.244	S24°38'11.10"	E033°09'98.20"	26
P6R	Marrambajane bridge	58.872	S24°41'30.30"	E033°15'65.50"	23

\*Elevation obtained by GPS placed at the top of the sampling station bridge.

farming [40]. The vitality of the scheme depends on two key hydraulic constructions, Macarretane Weir and Massingir Dam. These constructions are located upstream aiming at assuring regular functionality of 90% gravity-oriented transportation of sufficient water amount for rice, maize, and vegetables farmers. Surface irrigation systems dominates in the region, where furrow and flooding methods covers the most areas. Data was collected in three sets of stations established at each section of the principal canal, Upper sector (upstream section), South sector (or midstream section) and River sector (or downstream section), as indicated in the Figure 2.

Data collections were performed at the sampling sites for the Upper, South and River sectors. For Upper sector, Macarretane Weir Intake, Railways-Node and FIPAG bridges, constituted the most upstream sites. The South sector had Lionde, Massawasse and Conhane bridges sites, and the river sector had included Nico, Muianga and Marrambajane bridges for sediment sampling stations. An estimated distance of 60 km of main canal, in total, was covered in the study. Table 1 presents all sampling stations, coordinates and their cumulative distances from the main water Intake.

### Statistical Analysis

This work divided the physical characteristics of the canal and water into two main groups or factors, namely, canal channel factors (CCF) and water inflow factors (WIF). The multivariate statistical analysis (MSA) was therefore performed for CCF and WIF, for both dry (DS) and wet (WS) seasons. CCF considered variables related to sedimentation process, which included water flow (Q), water depth (WD), canal depth (CD), settling velocity (SV), critical shear stress (CSS), plasticity index (PI), electrical conductivity (EC), exchangeable sodium percentage (ESP), sodium adsorption ratio (SAR), sodium concentration (Na<sup>+</sup>), potassium concentration (K<sup>+</sup>) and combined concentration of calcium and magnesium (Ca<sup>2+</sup>+Mg<sup>2+</sup>). Whilst the WIF factors was related to water flow (Q), water depth (WD), water temperature (T), water velocity (WV), suspended sediment concentration (SSC), electrical conductivity (EC), total dissolved solids (TDS), turbidity (TURB),

water pH (pH), sodium concentration in water (Na<sup>+</sup>), potassium concentration in water (K<sup>+</sup>) and calcium concentration in water (Ca<sup>2+</sup>). The matrix of correlation together with the correlation coefficient (r) among variables were performed, for which the r ≥ 0.5 was significant. Thereafter, the factor loadings signifying the original variables contributions in the new factors, and the respective eigenvalues of each factor, produced by PCA, were obtained. Cluster analysis was also used to explain sampling sites relationship based on their similar sedimentation characteristics across CIS main canal. Lastly, the cross-section profiles and sediments flow were performed for the three hydraulic sectors of the CIS, namely, Montante (Upper sector), Sul (South sector) and Rio (River sector).

### Principal Component Analysis (PCA)

PCA consisting of data reduction statistical procedure was applied to lower dimensionality of the data set and identifying relevant spectral features holding significant influence of the spectral variance and to find hidden subtle relations within the referred data that contains different variables [41]. The choice of this technique was to reduce the dimension, models data, detects existing outliers, and allows selecting main variables, classifying them, validate, and run predictions of analyzed samples [29,42]. To achieve this, the PCA rotated the original variables contained in the data set and transformed it into newly generated variables projected through the new planes of maximum variance, where the first few axes do reflect the majority of the contained variations in the data. The newly generated variables became principal components (PCs), and these were normally orthogonal to each other, ensuring these new variables remain unrelated. Computation of the PCs were performed in order of maximum variance, signifying that PC 1 would then represent the largest variance amount, sequentially followed by PC 2 and PC 3.

The PCs were categorized in descendant order of importance as determined by their associated eigenvalues [43]. The technique was employed for exploratory analysis of data, this is, to search for underlying similarities or differences in the data set, and hence also referred to as unsupervised method. The common method for

determination of the optimum PCs number was through plotting the PCA eigenvalues, which in turn represents the PC variances, versus the PC number to obtain a scree plot [43]. The amount of the total variance of a principal component is generally represented by the eigenvalues [21,29,30,44,45]. The eigenvalue norm is then applied, which denotes that only the PCs with eigenvalues superior to one unit are considered as important [16,18].

The connection among the old and new variables is explained by the loadings plot. This plot fundamentally shows the influence of the old variables, such as for example, the water discharge in the irrigation system canal different depths. This permits regions of the spectra accounting for the grand part of variance, which constitutes the largest contribution, to be differentiated from noise, or simply the spectroscopic structures that are irrelevant to the PC's structure. All studied samples are assigned a specific score respecting to a particular PC, for which plotting two PCs one against the other, these samples become separated based on their scores per generated PC. The produced plot, also known as scores, plot, and consists of the method by which the elusive differences in spectra can be distinguished. Factor loadings indicate the variable contribution to the PC and how fairly the PC reflects the change of the same variable upon entire points in the data. Geometrically, factor loadings represent the angle cosine between the variable and the current PC. Thus, reduced the angle a greater the loadings. Therefore, factor loadings vary from -1 to +1.

### Cluster Analysis (CA)

CA applied for the CCF and WIF affecting sedimentation on the sampling sites located along the entire irrigation scheme, from the upstream to downstream. This analysis was performed establishing assemblages containing comparable sediments attributes, precisely for every station in the system. CA outputs appear in dendrogram, which is a depiction of the levels of connection among multivariate items [46-48]. The stations revealing the strongest resemblance were put closest all together [49-51]. All the statistical analyses both for PCA and CA were carried on the XLSTAT 2020.

## Results and Discussion

### Correlation Matrix

The PCA generates a correlation matrix that includes the correlation coefficient ( $r$ ) between all studied variables. For correlation coefficient  $r \geq 0.5$ , it was then considered significant. This study analysis was realized for canal channel factors (CCF) and water inflow factors (WIF), in two seasons, dry (DS) and wet (WS). CCF measured variables pertinent to processes of sedimentation, namely water flow (Q), water depth (WD), canal depth (CD), settling velocity (SV), critical shear stress (CSS), plasticity index (PI), electrical conductivity (EC), exchangeable sodium percentage (ESP), sodium adsorption ratio (SAR), sodium concentration ( $\text{Na}^+$ ), potassium concentration ( $\text{K}^+$ ) and combined concentration of calcium and magnesium ( $\text{Ca}^{2+}+\text{Mg}^{2+}$ ). For the CCF, in DS, a positive and high correlation coefficient is measured between CD and WD ( $r = 0.88$ ), PI and CSS ( $r = 0.96$ ), SAR and ESP ( $r = 0.80$ ),  $\text{Na}^+$  and SAR ( $r=0.78$ ) and  $\text{Ca}^{2+}+\text{Mg}^{2+}$  and PI ( $r = 0.73$ ).

A negative correlation coefficient was found for ESP and SV ( $r=-0.75$ ). These correlations reflect the physic-chemical characteristics of the main CIS canal during dry season. The fact that some of the physical variables appear with positive correlation sustains the theory that in somehow these variables do exert influence on each other and at some extent, to the chemical variables in the system, as per relation between  $\text{Ca}^{2+}+\text{Mg}^{2+}$  and PI. Moreover, sedimentation appears to be affected by canal channel factors in CIS. Besides this, it may also indicate anthropogenic factors affecting the system. The diagonalization of the correlation matrix was produced by PCA to avoid the problems of different measurement units of the original variables, since the standardization of all variables is automatically applied. Similar analysis was performed for the CCF during WS, where a positive and high correlation coefficient was found between CD and WD ( $r = 0.99$ ), WD and  $\text{Ca}^{2+}+\text{Mg}^{2+}$  ( $r=0.73$ ), CD and  $\text{Ca}^{2+}+\text{Mg}^{2+}$  ( $r=0.72$ ), CSS and PI ( $r=0.96$ ), CSS and EC ( $r=0.70$ ), PI and EC ( $r=0.78$ ), ESP and SAR ( $r=0.84$ ), ESP and  $\text{Na}^+$  ( $r=0.83$ ), SAR and  $\text{Na}^+$  ( $r=0.86$ ),  $\text{Na}^+$  and  $\text{Ca}^{2+}+\text{Mg}^{2+}$  ( $r=0.81$ ) and  $\text{K}^+$  and  $\text{Ca}^{2+}+\text{Mg}^{2+}$  ( $r=0.69$ ). A negative correlation coefficient was found for SV and CSS ( $r = -0.75$ ) and SV and PI ( $r=-0.69$ ). Tables 2 and 3, present Pearson correlation matrix ( $n$ ) for PCA on the canal channel factors (CCF) during DS and WS, respectively.

WIF covered water flow (Q), water depth (WD), water temperature (T), water velocity (WV), suspended sediment concentration (SSC), electrical conductivity (EC), total dissolved solids (TDS), turbidity (TURB), water pH (pH), sodium concentration in water ( $\text{Na}^+$ ), potassium concentration in water ( $\text{K}^+$ ) and calcium concentration in water ( $\text{Ca}^{2+}$ ). During DS, positive correlation coefficient was measured for Q and WV ( $r=0.98$ ), WD and EC ( $r=0.81$ ), WD and TDS ( $r=0.86$ ), EC and TDS ( $r=0.89$ ), pH and  $\text{Ca}^{2+}$  ( $r=0.91$ ), and  $\text{Na}^+$  and  $\text{K}^+$  ( $r=0.98$ ). No negative correlation coefficient is found in this season. But for WS, positive correlation coefficient was measured for Q and WV ( $r=0.92$ ), Q and TURB ( $r=0.85$ ), WD and  $\text{Na}^+$  ( $r=0.86$ ), T and SSC ( $r=0.96$ ), WV and TURB ( $r=0.90$ ), EC and TDS ( $r=0.99$ ), and  $\text{Ca}^{2+}$  and  $\text{K}^+$  ( $r=0.74$ ). Negative correlation coefficient is found for EC and TURB ( $r=-0.77$ ), and TDS and TURB ( $r=-0.74$ ). From these coefficients, one can see that sedimentation is affected by different inflow factors, in both seasons. Tables 4 and 5, details on the Pearson correlation matrix ( $n$ ) for PCA on the water inflow factors (WIF) for DS and WS.

### Factor Loadings

Factor loadings, expressing contributions from the authentic variables in the new factors, and the eigenvalues of each factor generated by PCA were calculated. The expanse of variance content explainable by given PC relies on the eigenvalue in relation to the cumulative eigenvalues. The number of PCs or factors to be selected in their identification process is grounded on various conditions to analyze data structure and comprehend the causal information [16]. The variance table (in %) of PCA factors, when a variance curve is given, a decreasing slope was observed from the sixth eigenvalue for the CCF and WIF during DS. For WS, the decreasing slope has been observed from the

**Table 2:** Pearson correlation matrix (n) for PCA on the canal channel factors (CCF) – during Dry Season.

Variables	Q	WD	CD	SV	CSS	PI	EC	ESP	SAR	Na <sup>+</sup>	K <sup>+</sup>	Ca <sup>2+</sup> + Mg <sup>2+</sup>
Q (m <sup>3</sup> /s)	<b>1</b>	0,416	0,313	-0,080	0,225	0,240	0,123	0,050	0,044	0,284	0,061	0,377
WD (m)	0,416	<b>1</b>	<b>0,879</b>	-0,217	0,348	0,284	0,286	0,014	-0,273	-0,065	-0,088	0,221
CD (m)	0,313	<b>0,879</b>	<b>1</b>	-0,279	0,337	0,279	0,155	0,114	-0,062	0,233	0,090	0,427
SV (m/s)	-0,080	-0,217	-0,279	<b>1</b>	-0,276	-0,216	0,095	<b>-0,753</b>	-0,585	-0,509	0,298	-0,104
CSS (N/m <sup>2</sup> )	0,225	0,348	0,337	-0,276	<b>1</b>	<b>0,958</b>	0,538	0,181	-0,238	0,169	-0,498	0,610
PI (%)	0,240	0,284	0,279	-0,216	<b>0,958</b>	<b>1</b>	0,514	0,124	-0,200	0,288	-0,415	<b>0,733</b>
EC (dS/m)	0,123	0,286	0,155	0,095	0,538	0,514	<b>1</b>	0,208	-0,042	0,031	-0,618	0,073
ESP (%)	0,050	0,014	0,114	<b>-0,753</b>	0,181	0,124	0,208	<b>1</b>	<b>0,799</b>	0,584	-0,623	-0,109
SAR (-)	0,044	-0,273	-0,062	-0,585	-0,238	-0,200	-0,042	<b>0,799</b>	<b>1</b>	<b>0,777</b>	-0,125	-0,068
Na <sup>+</sup> (mg/L)	0,284	-0,065	0,233	-0,509	0,169	0,288	0,031	0,584	<b>0,777</b>	<b>1</b>	0,048	0,571
K <sup>+</sup> (mg/L)	0,061	-0,088	0,090	0,298	-0,498	-0,415	-0,618	-0,623	-0,125	0,048	<b>1</b>	0,248
Ca <sup>2+</sup> + Mg <sup>2+</sup> (mg/L)	0,377	0,221	0,427	-0,104	0,610	<b>0,733</b>	0,073	-0,109	-0,068	0,571	0,248	<b>1</b>

**Table 3:** Pearson correlation matrix (n) for PCA on the canal channel factors (CCF) – during Wet Season.

Variables	Q	WD	CD	SV	CSS	PI	EC	ESP	SAR	Na <sup>+</sup>	K <sup>+</sup>	Ca <sup>2+</sup> + Mg <sup>2+</sup>
Q (m <sup>3</sup> /s)	<b>1</b>	0,368	0,329	-0,360	0,049	0,057	0,004	0,564	0,441	0,460	-0,178	0,246
WD (m)	0,368	<b>1</b>	<b>0,998</b>	-0,471	0,199	0,201	0,026	0,295	0,407	0,601	0,528	<b>0,728</b>
CD (m)	0,329	<b>0,998</b>	<b>1</b>	-0,471	0,214	0,216	0,039	0,236	0,364	0,563	0,558	<b>0,721</b>
SV (m/s)	-0,360	-0,471	-0,471	<b>1</b>	<b>-0,748</b>	<b>-0,688</b>	-0,501	-0,315	-0,518	-0,518	-0,507	-0,441
CSS (N/m <sup>2</sup> )	0,049	0,199	0,214	<b>-0,748</b>	<b>1</b>	<b>0,958</b>	<b>0,703</b>	0,140	0,408	0,340	0,522	0,228
PI (%)	0,057	0,201	0,216	<b>-0,688</b>	<b>0,958</b>	<b>1</b>	<b>0,780</b>	0,179	0,322	0,375	0,499	0,344
EC (dS/m)	0,004	0,026	0,039	-0,501	<b>0,703</b>	<b>0,780</b>	<b>1</b>	0,057	0,166	0,372	0,565	0,481
ESP (%)	0,564	0,295	0,236	-0,315	0,140	0,179	0,057	<b>1</b>	<b>0,841</b>	<b>0,828</b>	-0,108	0,422
SAR (-)	0,441	0,407	0,364	-0,518	0,408	0,322	0,166	<b>0,841</b>	<b>1</b>	<b>0,864</b>	0,265	0,434
Na <sup>+</sup> (mg/L)	0,460	0,601	0,563	-0,518	0,340	0,375	0,372	<b>0,828</b>	<b>0,864</b>	<b>1</b>	0,444	<b>0,809</b>
K <sup>+</sup> (mg/L)	-0,178	0,528	0,558	-0,507	0,522	0,499	0,565	-0,108	0,265	0,444	<b>1</b>	<b>0,685</b>
Ca <sup>2+</sup> + Mg <sup>2+</sup> (mg/L)	0,246	<b>0,728</b>	<b>0,721</b>	-0,441	0,228	0,344	0,481	0,422	0,434	<b>0,809</b>	<b>0,685</b>	<b>1</b>

Values showed in bold imply difference from 0 under significance level alpha = 0,05.

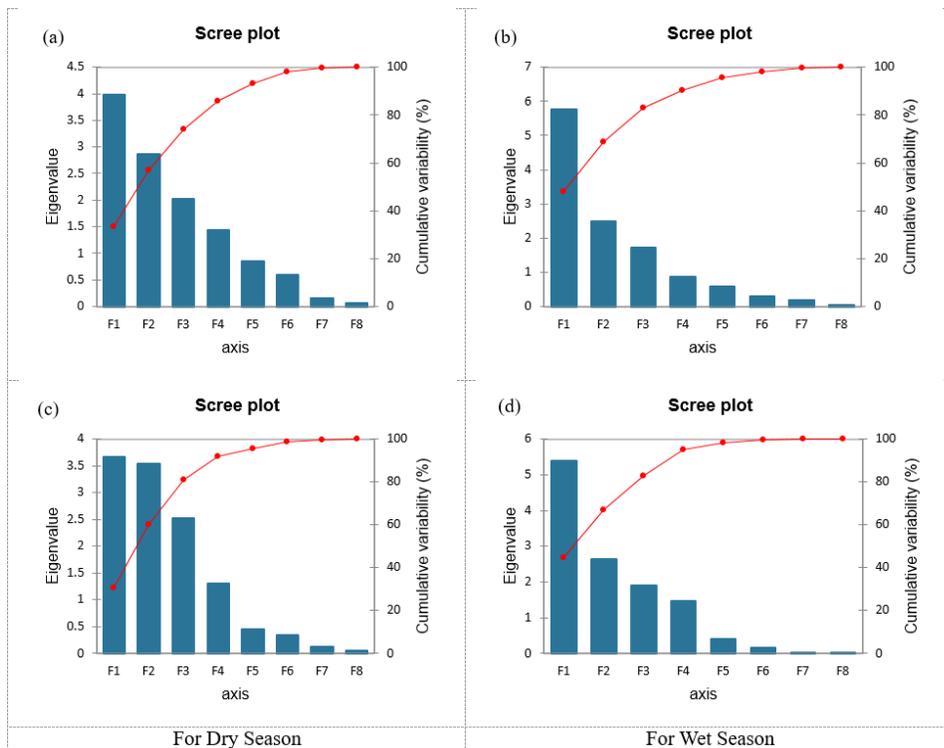
**Table 4:** Pearson correlation matrix (n) for PCA on the water inflow factors (WIF) – during Dry Season.

Variables	Q	WD	T	WV	SSC	EC	TDS	TURB	pH	Na <sup>+</sup>	K <sup>+</sup>	Ca <sup>2+</sup>
Q (m <sup>3</sup> /s)	<b>1</b>	0,416	-0,113	<b>0,984</b>	0,321	0,270	0,360	-0,541	0,102	-0,041	-0,014	0,100
WD (m)	0,416	<b>1</b>	-0,316	0,409	0,010	<b>0,811</b>	<b>0,863</b>	-0,252	-0,046	-0,191	-0,091	-0,016
T (°C)	-0,113	-0,316	<b>1</b>	-0,059	0,378	-0,627	-0,581	-0,364	-0,266	-0,486	-0,496	-0,099
WV (m/s)	<b>0,984</b>	0,409	-0,059	<b>1</b>	0,403	0,201	0,331	-0,623	0,132	-0,109	-0,098	0,146
SSC (mg/L)	0,321	0,010	0,378	0,403	<b>1</b>	-0,397	-0,241	-0,377	0,465	-0,152	-0,136	-0,422
EC (dS/m)	0,270	<b>0,811</b>	-0,627	0,201	-0,397	<b>1</b>	<b>0,885</b>	0,043	0,010	0,174	0,242	-0,044
TDS (mg/L)	0,360	<b>0,863</b>	-0,581	0,331	-0,241	<b>0,885</b>	<b>1</b>	0,097	0,160	0,204	0,249	-0,001
TURB (NTU)	-0,541	-0,252	-0,364	-0,623	-0,377	0,043	0,097	<b>1</b>	0,011	0,626	0,615	-0,307
pH (-)	0,102	-0,046	-0,266	0,132	-0,465	0,010	0,160	0,011	<b>1</b>	-0,252	-0,358	<b>0,905</b>
Na <sup>+</sup> (mg/L)	-0,041	-0,191	-0,486	-0,109	-0,152	0,174	0,204	0,626	-0,252	<b>1</b>	<b>0,982</b>	-0,551
K <sup>+</sup> (mg/L)	-0,014	-0,091	-0,496	-0,098	-0,136	0,242	0,249	0,615	-0,358	<b>0,982</b>	<b>1</b>	-0,625
Ca <sup>2+</sup> (mg/L)	0,100	-0,016	-0,099	0,146	-0,422	-0,044	-0,001	-0,307	<b>0,905</b>	-0,551	-0,625	<b>1</b>

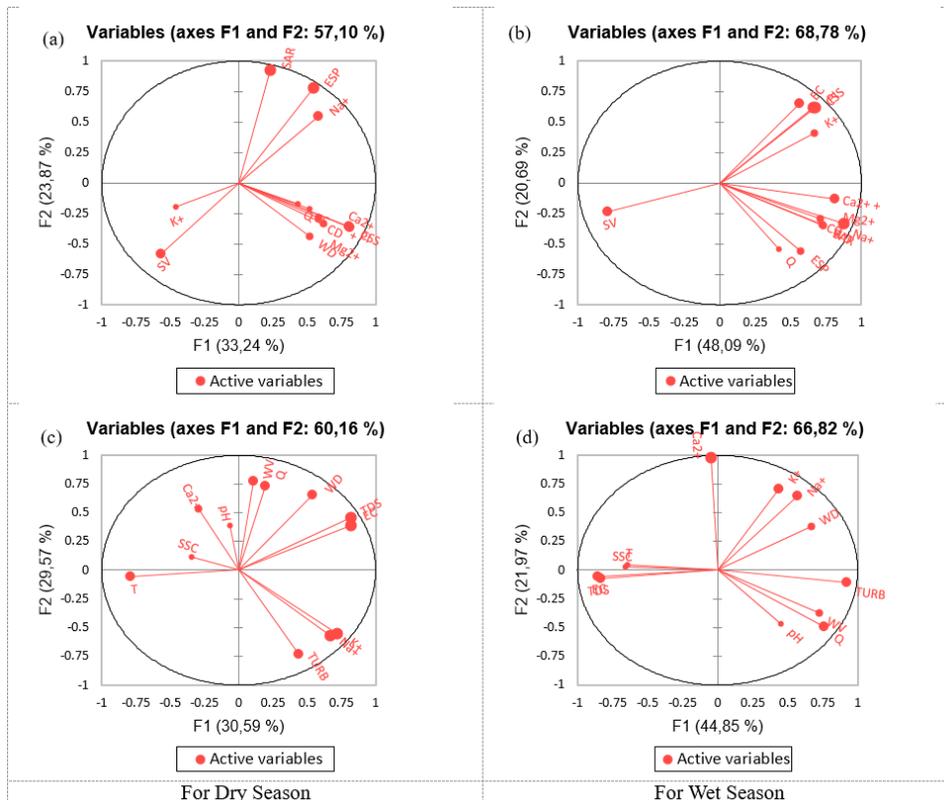
**Table 5:** Pearson correlation matrix (n) for PCA on the canal water inflow factors (WIF) – during Wet Season.

Variables	Q	WD	T	WV	SSC	EC	TDS	TURB	pH	Na <sup>+</sup>	K <sup>+</sup>	Ca <sup>2+</sup>
Q (m <sup>3</sup> /s)	<b>1</b>	0,368	-0,293	<b>0,918</b>	-0,361	-0,480	-0,450	<b>0,851</b>	0,592	0,163	0,110	-0,465
WD (m)	0,368	<b>1</b>	-0,427	0,290	-0,488	-0,319	-0,286	0,492	0,521	<b>0,860</b>	0,393	0,374
T (°C)	-0,293	-0,427	<b>1</b>	-0,070	<b>0,955</b>	0,640	0,648	-0,350	-0,152	-0,369	0,113	0,171
WV (m/s)	<b>0,918</b>	0,290	-0,070	<b>1</b>	-0,094	-0,523	-0,485	<b>0,896</b>	0,521	0,170	0,325	-0,334
SSC (mg/L)	-0,361	-0,488	<b>0,955</b>	-0,094	<b>1</b>	0,581	0,586	-0,390	-0,142	-0,363	0,068	0,134
EC (dS/m)	-0,480	-0,319	0,640	-0,523	0,581	<b>1</b>	<b>0,998</b>	<b>-0,773</b>	-0,146	-0,321	-0,469	0,015
TDS (mg/L)	-0,450	-0,286	0,648	-0,485	0,586	<b>0,998</b>	<b>1</b>	<b>-0,742</b>	-0,105	-0,302	-0,464	0,005
TURB (NTU)	<b>0,851</b>	0,492	-0,350	<b>0,896</b>	-0,390	<b>-0,773</b>	<b>-0,742</b>	<b>1</b>	0,414	0,417	0,510	-0,122
pH (-)	0,592	0,521	-0,152	0,521	-0,142	-0,146	-0,105	0,414	<b>1</b>	0,148	-0,195	-0,406
Na <sup>+</sup> (mg/L)	0,163	<b>0,860</b>	-0,369	0,170	-0,363	-0,321	-0,302	0,417	0,148	<b>1</b>	0,561	0,582
K <sup>+</sup> (mg/L)	0,110	0,393	0,113	0,325	0,068	-0,469	-0,464	0,510	-0,195	0,561	<b>1</b>	<b>0,737</b>
Ca <sup>2+</sup> (mg/L)	-0,465	0,374	0,171	-0,334	0,134	0,015	0,005	-0,122	-0,406	0,582	<b>0,737</b>	<b>1</b>

Values showed in bold imply difference from 0 under significance level alpha = 0,05.



**Figure 3:** PCA Eigenvalue plots for canal channel factors (a&b) and water inflow factors (c&d) for DS and WS from the variable plots for CCF, both components explained 57.10% for DS and 68.78% for WS. In DS (plot a), F1 is mostly associated with  $WD > CD > Ca^{2+} + Mg^{2+} > CSS > Q > EC$  and PI. In turn, F2 seemed very much associated with  $SAR > ESP$  and  $Na^+$ , negative association is found for  $K^+$  and SV. In WS (plot b), F1 appears very much associated with  $ESP > SAR > Na^+ > CD > Ca^{2+} + Mg^{2+}$  and Q, whilst F2 associated with  $CSS > PI > EC$  and  $K^+$ , and negatively associated with SV. This suggests that for both seasons, F1 is mostly comprised of canal depth, combination of calcium and magnesium and water flow, whilst for F2, no common positively related variables were found for both seasons, but a negative effect coming from the settling velocity.



**Figure 4:** PCA variable plots for canal channel factors (a&b) and water inflow factors (c&d) for DS and WS.

fourth eigenvalue. This trend is repeated for both groups of factors, CCF and WIF. Rendering to the percentage of total cumulative variance computed above 98.11% and 99.58%, were selected six factors during DS for CCF and WIF, respectively. Similarly, total cumulative variance above 90.51% and 95%, were selected four factors during WS for CCF and WIF, correspondingly. Factor F1 for CCF during DS, explained 33.24% of the variance and provided a positive correlation with variables such as WD, CD, CSS, PI, EC, ESP, Na<sup>+</sup>, and Ca<sup>2+</sup>+Mg<sup>2+</sup>, while correlating negatively with SV. Same for F1 of CCF, but for WS, it explained 48.09% variance and provided also positive correlation with variables WD, CD, CSS, PI, EC, ESP, SAR, Na<sup>+</sup>, K<sup>+</sup>, and Ca<sup>2+</sup>+Mg<sup>2+</sup>. This fact denotes high sedimentation potential through canal channel variables (CCF) of water depth, canal depth, shear stress, plasticity index, sodium related parameters. This may emphasize that canal channel containing sediment particles characterized by the presence of sodium, has less possibilities to aggregate and precipitate thereafter, but remains suspended towards downstream side of the canal. However, there are another part of flocculated sediment particles depositing in the canal bed while the fine ones move toward the scheme end [15,52-54].

WIF for F1 explained 30.59% and 44.85, for DS and WS, respectively. F1 in DS, appeared having positive correlation for variables of WD, EC, TDS, Na<sup>+</sup> and K<sup>+</sup>, whilst a negative correlation with T. During WS, positive correlation was detected for Q, WD, WV, TURB and Na<sup>+</sup>, while a negative correlation with T, WV, SSC, EC, and TDS. These outcomes suggest that sediment accumulating in the CIS are triggered by water flow, water depth, water velocity, turbidity and somehow the presence of sodium which impacts on the particle susceptibility to flocculate. Water temperature, water velocity, suspended sediment concentration, electric conductivity and total dissolved solids appear to be less pertinent in this factor. F2 of CCF, counted to explain 57.10% and 68.78% for DS and WS, respectively. During DS, a positive correlation with ESP, SAR, and Na<sup>+</sup> was found, while a negative correlation was found for SV. During WS, a positive correlation was found for CSS, PI, and EC, whilst a negative correlation for Q and ESP. F2 appears to indicate that concentration of sodium and its interaction to potassium and calcium are critically important for this component and less relevant to the settling velocity. F2 for WIF explained 60.16% for DS and 66.82% for WS. A positive correlation with Q, WD, WV and Ca<sup>2+</sup>, and negative correlation with TURB, Na<sup>+</sup> and K<sup>+</sup> was observed during dry season station. However, during WS, a positive correlation was observed for Na<sup>+</sup>, K<sup>+</sup>, and Ca<sup>2+</sup>. This seems to indicate that water discharge, water depth and velocity, and concentration of calcium strongly affect sedimentation, over water turbidity, calcium, and potassium. F3 of CCF explained however, 73.98% for DS and 83.19% for WS, whereby there was found a positive correlation with K<sup>+</sup>, and negative correlation for EC, in the dry season. Again, another positive correlation was found for WD and CD during wet season. WIF explained 81.16 % in DS and 82.72% in the WS.

A positive correlation was observed with for the SSC, and a negative correlation for the pH and Ca<sup>2+</sup>, in dry season, whilst during WS, no

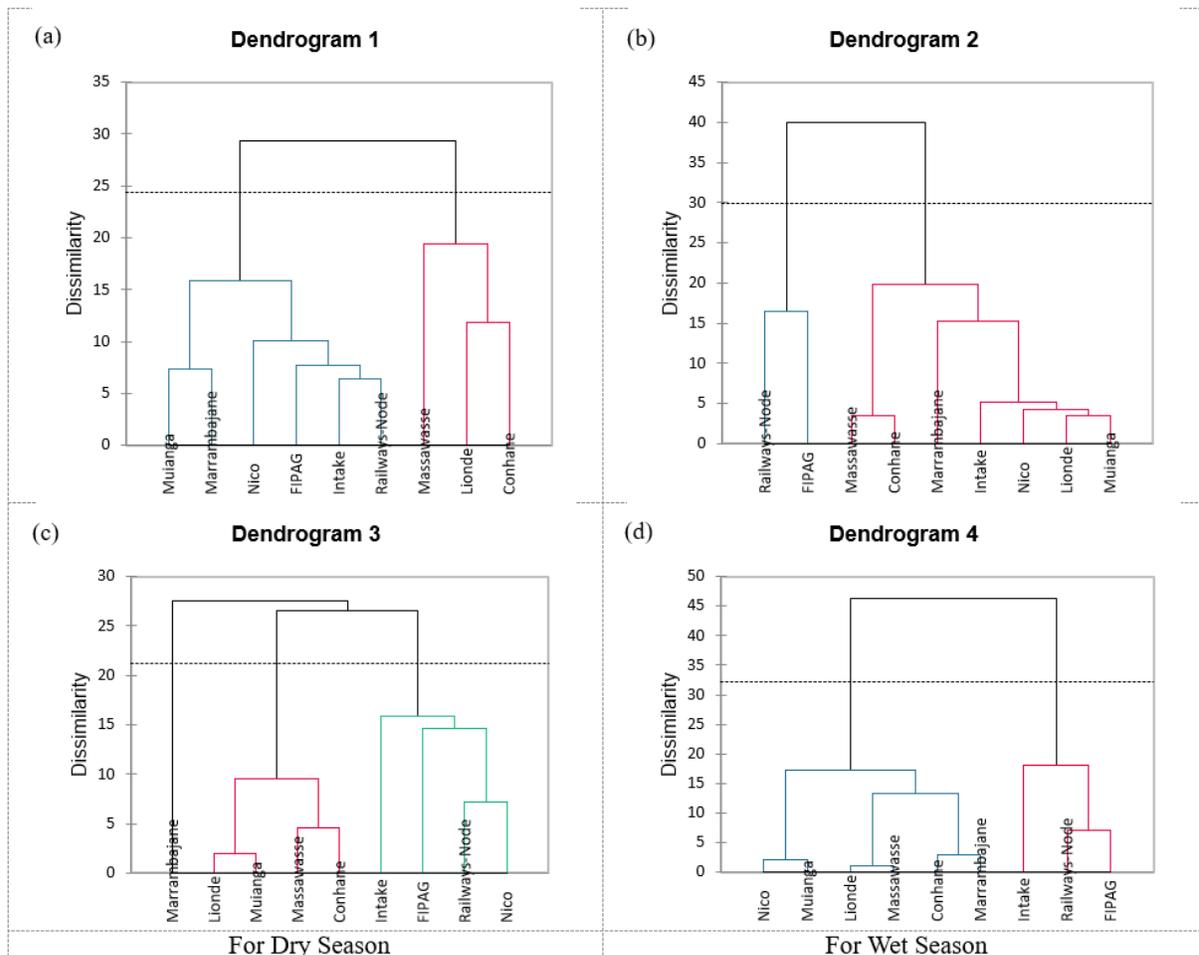
significant correlation was found at 95% of statistical significance level. DS for CCF presents significance for F4 and F5, like WS of CCF, and both seasons of WIF. For F4, a positive correlation was found for WD and CD, and F5 presented positive correlation with Q. Then, it appears that, these two factors, water depth and canal depth as well as, water discharge are seemingly important over any other variable in the irrigation system. The eigenvalue and respective cumulative (in %) values are presented in the scree plots in Figure 3, for CCF (a&b) and WIF (c&d), for DS and WS. From these figures one can see that the first three factors, F1, F2 and F3, account for above 80% of variance, except for the CCF in DS (Figure a), with 74%.

From inflow factors point of view, the components F1 and F2 explains in 60.16% and 66.82%, for DS and WS, respectively. For DS, F1 is well associated with Na<sup>+</sup> > K<sup>+</sup> and TURB, and negative association with Ca<sup>2+</sup> > SSC and pH. F2 is strongly associated with TDS > EC > WD > WV and Q, and negatively related to T. For WS, F1 has strong association coming from TURB > WV > Q and pH, while negative association with SSC and T. F2, is strongly associated with Na<sup>+</sup> > K<sup>+</sup> and WD, and negative association with EC and TDS. In general terms, for both seasons, strong associations between F1 and variables is found for turbidity and negative effect from suspended sediment concentration, meanwhile, for F2, positive association was found for water depth in the canal. Figure 4 presents variables plots for canal channel (a & b) and inflow (c & d) factors for dry and wet seasons.

### Cluster Analysis

From the cluster analysis a relationship between the sampling sites based on their similar sedimentation characteristics was found. The dendrograms in Figure 5 indicate that the sampling sites were grouped into two clusters (dendrograms 1, 2 and 4) and three clusters (dendrogram 3), both for CCF and WIF, for both seasons. For CCF, in DS, cluster 1 (C1) is comprised of Intake, Railways-Node, FIPAG, Nico, Muianga and Marrambajane sampling stations, and cluster 2 (C2) by Lionde, Massawasse and Conhane. However, for the wet season, C1 is composed of Railways-Node and FIPAG, whilst C2 made of Intake, Lionde, Massawasse, Conhane, Nico, Muianga and Marrambajane. Furthermore, for the WIF, during dry season, C1 is comprised only of Marrambajane station, C2 made of Lionde, Muianga, Massawasse and Conhane, and C3, include Intake, Railways-Node, FIPAG and Nico. Lastly, for the wet season, C1 is composed of Nico, Muianga, Lionde, Massawasse, Conhane and Marrambajane, and C2, made of Intake, Railways-Node and FIPAG. It is, therefore, worth reference that the cluster analysis has revealed stations that fall in the same cluster have common water flow and sediment parameters.

From the cluster analysis different hydraulic landscape features were shown for each site in terms of the physico-chemical characteristics. These features are of relevance to assessing the inflow sources, given the fact that the Intake station receives water coming directly from the Limpopo River mainstream, from which all system is then supplied. In addition, Railways-Node and Lionde



**Figure 5:** Cluster analysis for canal channel factors (a&b) and water inflow factors (c&d) for DS and WS.

stations present peculiarity by which are also diverting points for water to reach other downstream sites in the scheme. For this reason, these two stations present large chances for similarities in terms of water and sediment processes. Any potential difference may be due to the sampling seasons. However, when considering factor loadings and their respective variables associated, one can find that clusters agree with loadings, since the upstream stations face major amount of sediment inflow into the scheme, accumulating downstream. Changes in the water depths indicated higher levels at most upstream and lower at downstream, with opposite scenario for the sedimentation. Consequently, clusters appear to be always combined in accordance with the sites of occurrence, either if among the up-, mid- or down-stream.

### Conclusions

This study identified key factors influencing sedimentation at Chókwe Irrigation Scheme principal canal, with application of two multivariate statistical procedures, specifically, principal component analysis and cluster analyses. Results suggest that F1 on CCF in DS, explained 33.24% of the variance and positively correlated with water depth (WD), canal depth (CD), critical shear stress (CSS), plasticity index (PI), electrical conductivity (EC), exchangeable sodium percentage (ESP), sodium concentration

( $\text{Na}^+$ ), and combined concentration of calcium and magnesium ( $\text{Ca}^{2+}+\text{Mg}^{2+}$ ) and correlated negatively with settling velocity (SV). During WS, F1 enlightened 48.09% variance and correlated positively with WD, CD, CSS, PI, EC, ESP, sodium adsorption ratio (SAR),  $\text{Na}^+$ , potassium concentration ( $\text{K}^+$ ), and  $\text{Ca}^{2+}+\text{Mg}^{2+}$  variables. On the other hand, for WIF, F1 explained 30.59% for DS and 44.85 for WS. F1 during dry season, correlated positively with WD, EC, total dissolved solids (TDS),  $\text{Na}^+$  and  $\text{K}^+$  variables, whilst a negative correlation was found with water temperature (T). These findings on CCF suggest that sedimentation tend to occur more often during WS than in DS. Similarly, it appears that WIF has more contribution for sedimentation than the CCF do. Suggestively, the cluster analysis pointed to a difference in hydraulic landscape features at the sites, with respect to the physicochemical characteristics. Cluster analysis also allowed grouping sampling sites grounded on their degree of parallel within a class or distinctions between classes. Sampling stations located in the Upstream sites (Montante sector) appeared more closely related among each other than with other stations. Midstream sites (Sul sector) and downstream sites (Rio sector) sampling stations presented close related pattern in terms of sedimentation. The study revealed that canal channel factors (CCF) and water inflow factors (WIF) are behind sedimentation occurrence at CIS, with

geometric, water inflow-flux and physicochemical parameters explaining the causes. It is, therefore, recommended that these factors be paid attention to for irrigation system management and canal maintenance process. Differences on the type of factors analyzed did not provide enough evidence to ascertain which one plays the most role for sedimentation. This remains undisclosed under the current conditions in the scheme. Therefore, future research work on determining the specific weight of physical and chemical factors on the sedimentation may contribute to better explaining CIS hydraulic processes.

### Acknowledgments

The authors express their gratitude to Instituto Superior Politécnico de Gaza (ISPG) and the Netherlands Initiative for Capacity Development in Higher Education (NICHE) for the NICHE/MOZ/150 project that provided the research funding. Additionally, authors thank the Hidráulica de Chókwè, Empresa Pública (HICEP) for allowing access to the Chókwè Irrigation Scheme and to the relevant historical data set. Special acknowledgement to Q-Point, B.V., for encouragement and support in this study.

### References

1. Rath A, Swain PC. Evaluation of performance of irrigation canals using benchmarking techniques –a case study of Hirakud dam canal system, Odisha, India. *ISH J. Hydraul. Eng.* 2020; 26: 51-58.
2. Terefe Y, Singh P. East-bank canal water delivery performance evaluation: Case study of Finchaa Sugar Estate, Ethiopia. *ISH J. Hydraul. Eng.* 2020; 1-9.
3. Kuriachen P, Suresh A, Aditya KS, et al. Irrigation development and equity implications: The case of India. *Int. J. Water Resour. Dev.* 2021: 1-17.
4. Goes BJM, Clark AK, Bashir K. Water allocation strategies for meeting dry-season water requirements for Ganges Kobadak Irrigation Project in Bangladesh. *Int. J. Water Resour. Dev.* 2021; 37: 300-320.
5. Rath A, Swain PC. Optimal allocation of agricultural land for crop planning in Hirakud canal command area using swarm intelligence techniques. *ISH J. Hydraul. Eng.* 2021; 27: 38-50.
6. Ricart S, Rico A, Kirk N, et al. How to improve water governance in multifunctional irrigation systems? Balancing stakeholder engagement in hydrosocial territories. *Int. J. Water Resour. Dev.* 2019; 35: 491-524.
7. Yalcin E. Estimation of irrigation return flow on monthly time resolution using SWAT model under limited data availability. *Hydrological Sciences Journal*, 2019.
8. Zhang G, Tang L, Liu Z, et al. Machine-learning-based damage identification methods with features derived from moving principal component analysis. *Mechanics of Advanced Materials and Structures.* 2020; 27: 1789-1802.
9. Ettema R, Armstrong DL. Bedload and channel morphology along a braided, sand-bed channel: Insights from a large flume. *J. Hydraul. Res.* 2019; 57: 822-835.
10. Montes C, Berardi L, Kapelan Z, et al. Predicting bedload sediment transport of non-cohesive material in sewer pipes using evolutionary polynomial regression – multi-objective genetic algorithm strategy. *Urban Water Journal.* 2020; 17: 154-162.
11. Schwindt S, Franca MJ, Schleiss AJ. Bottom slope influence on flow and bedload transfer through contractions. *J. Hydraul. Res.* 2019; 57: 197-210.
12. Sharma A, Kumar B. Flow resistance in seepage-affected alluvial channel. *ISH J. Hydraul. Eng.* 2020; 26: 127-137.
13. Singh PK, Khatua KK, Banerjee S. Flow resistance in straight gravel bed inbank flow with analytical solution for velocity and boundary shear distribution. *ISH Journal of Hydraulic Engineering.* 2018: 1-14.
14. Bouguerne A, Boudoukha A, Benkhaled A, et al. Assessment of surface water quality of Ain Zada dam (Algeria) using multivariate statistical techniques. *Int. J. River Basin Manag.* 2017; 15: 133-143.
15. Croux C, García-Escudero L A, Gordaliza A, et al. Robust principal component analysis based on trimming around affine subspaces. *Statistica Sinica.* 2017; 27: 1437-1459. <https://www.jstor.org/stable/26383263>
16. Gyimah RAA, Gyamfi C, Anornu GK, et al. Multivariate statistical analysis of water quality of the Densu River, Ghana. *J. River Basin Manag.* 2021: 1-11.
17. Nguyen HD, Quan NH, Quang NX, et al. Spatio-temporal pattern of water quality in the Saigon-Dong Nai river system due to wastewater pollution sources. *J. River Basin Manag.* 2019:1-23.
18. Souza FGde, Campos MCC, Pinheiro EN, et al. Aggregate stability and carbon stocks in Forest conversion to different cropping systems in Southern Amazonas, Brazil. *Carbon Management.* 2020; 11: 81-96.
19. Bostanmaneshrad F, Partani S, Noori R, et al. Relationship between water quality and macro-scale parameters (land use, erosion, geology, and population density) in the Siminehrood River Basin. *Sci Total Environ.* 2018; 639: 1588-1600.
20. Ait-Sahalia Y, Xiu D. Principal Component Analysis of High-Frequency Data. *Journal of the American Statistical Association.* 2019; 114: 287-303.
21. Landgraf AJ, Lee Y. Generalized Principal Component Analysis: Projection of Saturated Model Parameters. *Technometrics.* 2020; 62: 459-472.
22. Fan W, Zhou J, Zhou Y, et al. Heavy metal pollution and health risk assessment of agricultural land in the Southern Margin of Tarim Basin in Xinjiang, China. *Int J Environ Health Res.* 2021; 31: 835-847.
23. Hui T, Jizhong D, Shimin M, et al. Application of water quality index and multivariate statistical analysis in the hydrogeochemical assessment of shallow groundwater in Hailun, northeast China. *Human and Ecological Risk Assessment: An International Journal.* 2021; 27: 651-667.

24. Kabir MdH, Tusher TR, Hossain MdS, et al. Evaluation of spatio-temporal variations in water quality and suitability of an ecologically critical urban river employing water quality index and multivariate statistical approaches: A study on Shitalakhya river, Bangladesh. *Human and Ecological Risk Assessment: An International Journal*. 2021; 27: 1388-1415.
25. Li W, Xie X, Song J, et al. Assessment and source identification of toxic metals in an abandoned synthetic diamond production plant from Anhui Province, China. *Environ. Forensics*. 2021; 22: 340-350.
26. Liu JP, DeMaster DJ, Nguyen TT, et al. Stratigraphic Formation of the Mekong River Delta and Its Recent Shoreline Changes. *Oceanography*. 2017; 30: 72-83. <http://www.jstor.org/stable/26201900>
27. Prajapati M, Jariwala N, Agnihotri P. Evaluation of groundwater quality with special emphasis on fluoride contamination using multivariate statistical analysis in rural parts of Surat district, Gujarat. *ISH J. Hydraul. Eng.* 2020; 26: 179-186.
28. Khan MB, Dai X, Ni Q, et al. Toxic Metal Pollution and Ecological Risk Assessment in Sediments of Water Reservoirs in Southeast China. *Soil Sediment Contam.* 2019; 28: 695-715.
29. Shen Y, Xu H, Liu X. An alternating minimization method for robust principal component analysis. *Optimization Methods and Software*. 2019; 34: 1251-1276.
30. Zakhem B A, Al-Charidch A, Kattaa B. Using principal component analysis in the investigation of groundwater hydrochemistry of Upper Jezireh Basin, Syria. *Hydrol. Sci. J.* 2017; 62: 2266-2279.
31. Zeinalzadeh K, Rezaei E. Determining spatial and temporal changes of surface water quality using principal component analysis. *J. Hydrol. Reg. Stud.* 2017; 13: 1-10.
32. Gaagai A, Boudoukha A, Boumezbeur A, et al. Hydrochemical characterization of surface water in the Babar watershed (Algeria) using environmetric techniques and time series analysis. *J. River Basin Manag.* 2017; 15: 361-372.
33. Mathew J, Kumar RV. Multilinear Principal Component Analysis with SVM for Disease Diagnosis on Big Data. *IETE Journal of Research*. 2019: 1-15.
34. Simmatis B, Nelligan C, Rühländ KM, et al. Tracking 200 years of water quality in Muskrat Lake, a eutrophic lake trout lake in Ontario (Canada) with cyanobacterial blooms. *Lake and Reservoir Management*. 2020; 36: 260-277.
35. El Hawary A, Shaban M. Improving drainage water quality: Constructed wetlands-performance assessment using multivariate and cost analysis. *Water Science*. 2018; 32: 301-317.
36. Sepehri M, Ghahramani A, Kiani-Harchegani M, et al. Assessment of drainage network analysis methods to rank sediment yield hotspots. *Hydrological Sciences Journal*. 2021; 66: 904-918.
37. Tonini M, Pecoraro G, Romailier K, et al. Spatio-temporal cluster analysis of recent Italian landslides. *Georisk: Assessment and Management of Risk for Engineered Systems and Geohazards*. 2020; 16: 1-19.
38. Agarwal A, Xue L. Model-Based Clustering of Nonparametric Weighted Networks with Application to Water Pollution Analysis. *Technometrics*. 2020; 62: 161-172.
39. De Sousa LS, Wambua RM, Raude JM, et al. Cohesive and non-uniform sediment characteristics in the unlined canal of Chókwè Irrigation Scheme, Mozambique. *J. Sediment. Environ.* 2021.
40. De Sousa LS, Wambua RM, Raude J M, et al. Assessment of Water Flow and Sedimentation Processes in Irrigation Schemes for Decision-Support Tool Development: A Case Review for the Chókwè Irrigation Scheme, Mozambique. *AgriEngineering*. 2019; 1: 100-118.
41. Varmuza K, Filzmoser P. Introduction to multivariate statistical analysis in chemometrics. Boca Raton, CRC Press. [PDF] Introduction to Multivariate Statistical Analysis in Chemometrics | Semantic Scholar. 2009.
42. Miao C, Lv Z. Nonlinear chemical process's fault detection based on adaptive kernel principal component analysis. *Systems Science & Control Engineering*. 2020; 8: 350-358.
43. Setiawan O, Sartohadi J, Hadi MP, et al. Infiltration characterization using principal component analysis and K-means cluster analysis on quaternary volcanic landscape at the southern flank of Rinjani Volcano, Lombok Island, Indonesia. *Physical Geography*. 2020; 41: 217-237.
44. Wang Y, Aelst SV. Sparse Principal Component Analysis Based on Least Trimmed Squares. *Technometrics*. 2020; 62: 473-485.
45. Zhang Y, Xiong Y, Chao Y, et al. Hydrogeochemistry and quality assessment of groundwater in Jinghui canal irrigation district of China. *Human and Ecological Risk Assessment: An International Journal*. 2020; 26: 2349-2366.
46. Peixoto HJC, Baia LP, Pereira DR, et al. Sedimentological Sectorization Model In An Amazonian Estuary. *Journal of Coastal Research*. 2018; 91-95.
47. Scrucca L, Serafini A. Projection Pursuit Based on Gaussian Mixtures and Evolutionary Algorithms. *J Comput Graph Stat.* 2019; 28: 847-860.
48. Sun K, Rajabtabar M, Samadi S, Rezaie-Balf M, et al. An integrated machine learning, noise suppression, and population-based algorithm to improve total dissolved solids prediction. *Eng. Appl. Comput. Fluid Mech.* 2021; 15: 251-271.
49. Morgan GB. Generating nonnormal distributions via Gaussian mixture models. *Structural Equation Modeling: A Multidisciplinary Journal*. 2020: 1-11.
50. Sutliff-Johansson S, Pontér S, Mäki A, et al. Groundwater environmental forensic investigation combining multivariate statistical techniques and screening analyses. *Environmental Forensics*. 2020: 1-15.

- 
51. Zamprogno B, Reisen V A, Bondon P, et al. Principal component analysis with autocorrelated data. *J Stat Comput Simul.* 2021; 1-19.
  52. Kansara S, Parashar S, Xue Z. Effective Decision Making and Data Visualization Using Partitive Clustering and Principal Component Analysis (PCA) for High Dimensional Pareto Frontier Data. *SAE Int. J Mater Manuf.* 2015; 8: 336-343. <https://www.jstor.org/stable/26268720>
  53. Rokni K, Ahmad A, Solaimani K, et al. A New Approach for Detection of Surface Water Changes Based on Principal Component Analysis of Multitemporal Normalized Difference Water Index. *Journal of Coastal Research.* 2016; 32: 443-451. <https://www.jstor.org/stable/43752206>
  54. See JCG, Porio EE. Assessing Social Vulnerability to Flooding in Metro Manila Using Principal Component Analysis. *Philippine Sociological Review.* 2015; 63: 53-80. <https://www.jstor.org/stable/24717187>