Original Paper   http://ajol.info/index.php/ijbcs   http://indexmedicus.afro.who.int

Application of multivariate statistical methods for the assessment of the variability of on-site sanitation faecal sludge in Cameroon

Wilfried Arsène LETAH NZOUEBET[1], Ebenezer SOH KENGNE[1],
Guy Valerie DJUMYOM WAFO[1], Alex Joel NGUENANG NJIOJOB[1],
Christoph HÖSER[2], Pierre NBENDAH[1], Andrea RECHENBURG[2] and
Ives Magloire KENGNE NOUMSI[1*]

[1]*Wastewater Research Unit, Faculty of Science, University of YaoundeI, P.O. Box 812 Yaounde, Cameroon.*
[2]*GeoHealth Centre, Institute for Hygiene and Public Health, University of Bonn, Sigmund Freud Str. 25, D. 53105 Bonn, Germany.*
[*]*Corresponding author; E-mail: ives_kengne@yahoo.fr*

ABSTRACT

This study examines quality and variability of FS collected in Yaounde City. Hierarchical cluster analysis (HCA) and principal component analysis (PCA) were applied for the description of FS data obtained in the context of Yaounde (Cameroon). Almost 13 parameters for the appreciation of the FS characteristics (temperature, pH, electric conductivity, salinity, COD, BOD_5, NH_4^+, TKN, TS, TVS, TSS, water content and total helminth eggs) were measured in 53 different samples originated from various on-sanitation systems found around the town represented by septic tanks, traditional latrines, ventilated improved pit latrines and piped equipped latrines located all around the town. FS samples were classified into 9 clusters with 72.91% of variation within the samples of the same cluster and 27.09% of variation between clusters. PCA of the whole data set indicated about 78% of the total variance with the first component accounting for 28.38% of the total variance and significantly correlated with COD, BOD_5, NH_4^+, TKN, TS and total helminth eggs. The second component accounting for 18.54% of the total variance correlated with electric conductivity, salinity and TVS. We found that the quality of FS is significantly different ($p< 0.05$) between groups of latrines. The *p*-values obtained after the Kruskall Wallis test were 0.03, 0.02, 0.01, 0.05 and 0.002 respectively for the parameters salinity, TKN, DCO, TSS and Water content. On the base of this study, it can be concluded that PCA and HCA could be helpful for the representation and interpretation of high variable FS quality data produced in urban and rural area of developing countries.
© 2019 International Formulae Group. All rights reserved

Keywords: Faecal sludge quality, descriptive analysis, hierarchical cluster analysis, principal component analysis, on-site sanitation, developing countries.

INTRODUCTION

Faecal sludge (FS) refers to the raw or partially digested slurry or solid that results from the storage of black water or excreta in on-site sanitation systems (OSS) (Tilley et al., 2014; Anh et al., 2018) The sanitation needs of 2.7 billion people worldwide are served by on-site sanitation technologies and that number is expected to grow to 5 billion by 2030 (Strande et al., 2014). In many cities of developing countries, on-site technologies have much wider coverage than sewer

systems. In Sub-Saharan Africa, 65-100% of sanitation access in urban areas is provided through on-site technologies represented by pit latrines, public latrines, septic tanks, aqua privies and dry toilets (Koné and Strauss, 2004; Cofie et al., 2006; Niwagaba et al., 2014; Kengne et al.; 2008; Bassan et al., 2013; Koné et al., 2016). These on-site sanitation systems required time to time removal and huge amount of FS is discharge daily into the environment without treatment. The most important step in designing the treatment technologies is to quantify and characterize FS to be treated. Unfortunately, the data on FS characteristics are still lacking in the world (Strande et al., 2014). Additionally, FS quality is difficult to categorize due to the variety of on-site technologies in use and the environmental conditions affecting its quality. Indeed, several environmental factors have been shown to affect the quality and variability of FS (Strauss et al., 1997; Strauss and Montangero, 2002; Fidjeland et al., 2013). These factors include among others the storage duration in on-site sanitation facilities, intrusion of groundwater, toilet usages and climate, and the population feeding habits. A simplified approach of FS classification based on the on-site sanitation sources have been proposed by Heinss et al. (1998). The authors divided faecal sludge into two classes, the low strength sludge which are most stabilized (septage) and the high strength sludge highly concentrate coming from public toilets and bucket latrines. However, this classification of Heinss et al. (1998) may be limited due to the overlapping of FS quality data that may occur and could influence on the designing of treatment systems. Furthermore, FS characteristic is difficult to determine due to the variety of households OSS in use. The classification should also take into consideration the whole FS quality data independent to the origin of samples. Green (2014) and Chatfied and Collins (2013) demonstrated that the multivariate statistical techniques are appropriate tools for a meaningful data interpretation. In this study, two different multivariate statistical methods, the principal component analysis (PCA) and hierarchical cluster analysis (HCA) as has

been used to assess a data set of FS characteristics from various OSS present of Yaounde.

MATERIALS AND METHODS
Context of the study area and faecal sludge management

The present study was carried out in Yaounde, the Capital City of Cameroon. Yaounde is an urban area of approximately 256 km^2, limited by latitudes 03°45' and 04° N and longitudes 11°20' and 11°40' E. The population of Yaounde was estimated at about 2.4 million inhabitants in 2011 (BUCREP, 2012). Yaounde like many other urban centres in developing countries is facing the problem of overpopulation and a large part of the capital consists of slums with very basic water supply, sanitation and waste disposal infrastructure (Parrot et al., 2009). Yaounde has an equatorial climate with four seasons comprising a long dry season (December-February), short rainy season (March-June), short dry season (July-August) and a long rainy season (September-November) (Lienou et al., 2008). The average annual rainfall is 1600 mm with an average temperature of 23 °C. The faecal sludge management in Yaounde is dramatic. The town operated without FS treatment plan and it was estimated at about 730 to 900 m^3 of FS discharged weekly in surrounding area of the town (Nomayos) without treatment thereby exposing the population of the town to faecal-oral transmitted diseases (Berteigne, 2012). In a survey conducted in some main quarters of Yaounde, it was found that approximately 75% of the FS were evacuated through traditional pit latrines, 20% through septic tanks and less than 2% in ventilated improved pit latrine (Ambassa Ndiomo, 2005). Kengne et al. (2009) mentioned that due to the dense hydrographic network of Yaounde, large quantities of FS are still disorderly discharged into drains and streams when the facilities are full, practice in disregard with the environment protection and sustainable development rules. For this study, investigations were carried out in four types of on-site sanitation systems found in Yaounde represented by traditional pit latrines, septic

tanks, VIP (ventilated improved pit) latrines and piped equipped latrines (Figure 1).

Traditional latrine is one of the most widely used on-site sanitation technologies in the study area where excreta along with anal cleansing materials (water or solids) are disposed into the pit. Septic tank is a watertight chamber made of concrete, fiberglass, PVC or plastic, through which black water flows for primary treatment. VIP latrine is an improved traditional latrine because continuous airflow through the ventilation pipe vents odours and acts as a trap for flies as they escape towards the light. Piped equipped latrine is one in which the top of the pit is relate to a PVC pipe allowing the exit of faecal sludge in case of flooding conditions. It is one of the latrines found in the study area principally located in the lowland area.

Monitored parameters and analytical methods
Faecal sludge collection methods

A total of 53 samples belonging to 30 pit latrines, 13 septic tanks, 04 VIP latrines and 6 piped equipped latrines were collected in various location of the town (Figure 1). The sampling material for this study was made up to an assembly comprising: a metal bar (1.5 m of length, 3 cm of diameter and a weight of 7 kg), fixed to a metallic box (10 cm in diameter and 20 cm height) welded at its lower extremity for the sludge sample collection; the upper end was connected to a 15 m graduated rope (Figure 2).

The use of the rope was made to circumvent difficulties which may occur due to deeper latrines usually found in the study area. The high mass of the metal bar allowed its easy entrance into the pit substrate. The sampling method took into account a number of considerations including; a non-destructive operation i.e the removal of sludge through the hole of the latrine slabs and easy cleaning of sampling material and latrine after the sampling operations. Sampling was performed in the first meter of sludge thickness after vigorous stirring. The precise sampling depth was at the discretion of the field (sampling) team, according to the observations made on

the devices about the actual depth of the toilet (height between the surface of the sludge layer and the slab of latrine, sometimes less than 2 m). The sampling protocol aimed at the introduction of sampling equipment through the hole of the latrine slabs; stirring (2 to 3 minutes) the sampling zone; extraction of a two or three fractions of sample (about 0.6 L); discharge into a 5 L bucket, stirring, transfer into sterilized plastic box and transportation of samples to the laboratory in refrigerated thermic boxes to avoid majors microbial and chemical changes.

Analytical methods

Analysis of FS included parameters such as temperature (Temp), electrical conductivities (EC), salinity, biological oxygen demand for five days (BOD_5), chemical oxygen demand (COD), the total Kjeldahl nitrogen (TKN), nitrogen ammonia (NH_4^+), total volatile solids (TVS), total solids (TS), total suspended solids (TSS), and water content. These parameters were measured following standard protocols for water and wastewater analyses (APHA/AWWA/WEF, 2005). The determination of the total helminth eggs (HE) in samples was done as described by Bastos et al. (2013). The descriptive statistic for FS samples is summarized in the Table 1.

Data treatment and multivariate statistical methods
Descriptive analysis of data

The descriptive statistical analysis of the physico-chemical and parasitic data of FS samples were done using the statistic software R, version 3.2.2 (Hothorn and Everitt, 2014) and the software XL Statistic. The obtained values were expressed in terms of minimum, maximum, mean, standard error of the mean (SE), standard deviation (Sd), coefficient of variation (CV), interquartile range (IQR) and skewness (Table 1). The level of variation was observed at 95% of confidence interval. The cross correlation matrix between some analysed parameters (BOD_5, COD, NH_4^+, TKN) was also performed to show the variation in the distribution of parameters in function of the type of latrines. Due to the

non-normal distribution in the number of samples within latrine groups and the values of the 13 parameters analysed, the non-parametric Kruskal-Wallis test (Acar and Sun, 2013) was used to determine the statistical differences between the latrine groups investigated. The Group of latrines of sizes n_j, j=1...m (Eq. 1), were combined and ranked in ascending order of magnitude. For this type of methodology, the Ties values (repeated value) are assigned to the average ranks. The n denotes the overall sample size, and the sum of ranks denoting for the jth sample is expressed by the following equation (Acar and Sun, 2013):

$$R_j = \sum_{i=1}^{nj} R(X_{ji}) \text{ (Eq. 2)}$$

The Kruskall-Wallis one-way analysis-of-variance test, H is defined as:

$$H = \frac{1}{S^2} \left\{ \sum_{j=1}^{m} \frac{R_j^2}{n_j} - \frac{n(n+1)2}{4} \right\} \text{ (Eq. 3)}$$

Where

$$S^2 = \frac{1}{n-1} \left\{ \sum_{all\ ranks} R(X_{ji})^2 - \frac{n(n+1)^2}{4} \right\}$$
(Eq. 4)

In this study there was no ties value (i.e. two or more values were not the same) and the above equation (3) was simplified to:

$$H = \frac{12}{n(n+1)} \sum_{j=1}^{m} \frac{R_j^2}{n_j} - 3(n+1) \text{ (Eq. 5)}$$

The sampling distribution of H was approximately X^2 with m-1 degrees of freedom.

Hierarchical Cluster analysis (HCA)

HCA is an unsupervised pattern recognition technique that uncovers intrinsic structure or underlying behaviour of a data set without making a priori assumption about the data, in order to classify the objects of the system into categories or clusters based on their nearness or similarity (Anderberg, 2014). Hierarchical clustering is the most common approach in which clusters are formed sequentially, by starting with the most similar pair of objects and forming higher clusters step by step. The Euclidean distance usually gives the similarity or dissimilarities between two samples and a distance can be represented by the difference between analytical values from both samples (Patras et al., 2011). HCA was performed in this study by using the

Ward's method: The Euclidean distances between centroid objects of classes was used to measure the similarity between classes of FS. This method uses the analysis of variance approach to evaluate the distances between clusters, attempting to minimize the sum of squares of any two clusters that can be formed at each step (Singh et al., 2013 ; Shrestha and Kazama, 2007). The linkage distance between classes (Height) is represented on y-axis. Due to the management of space, the linkage distance between the classes of FS samples in this paper is represented on x-axis.

PCA techniques extract the eigenvalues and eigenvectors from the covariance matrix of original variables. The PCs are the uncorrelated (orthogonal) variables, obtained by multiplying the original correlated variables with the eigenvector, which is a list of coefficients (loadings or weightings) (Liu et al., 2003). Thus, the PCs are weighted linear combinations of the original variables. PC provides information on the most meaningful parameters, which describe whole data set affording data reduction with minimum loss of original information (Liu et al., 2015). It is a powerful technique for pattern recognition that attempts to explain the variance of a large set of inter-correlated variables and transforming into a smaller set of independent (uncorrelated) variables (principal components). The principal component can be expressed as (Liu et al., 2015):

$$z_{ij} = a_{i1}x_{1j} + a_{i2}x_{2j} + a_{i3}x_{3j} + \dots + a_{im}x_{mj}$$
(Eq. 6)

Where z is the component score, a the component loading, x the measured value of variable, i the component number, j the variable and m the total number of variables. PCA of the FS data set was performed to extract significant PCs and to further reduce the contribution of variables with minor significance; these PCs were subjected to varimax rotation (raw) generating varifactors (VFs) (Shrestha and Kazama, 2007). As a result, a small number of factors will usually account for approximately the same amount of information as do the much larger set of original observation.
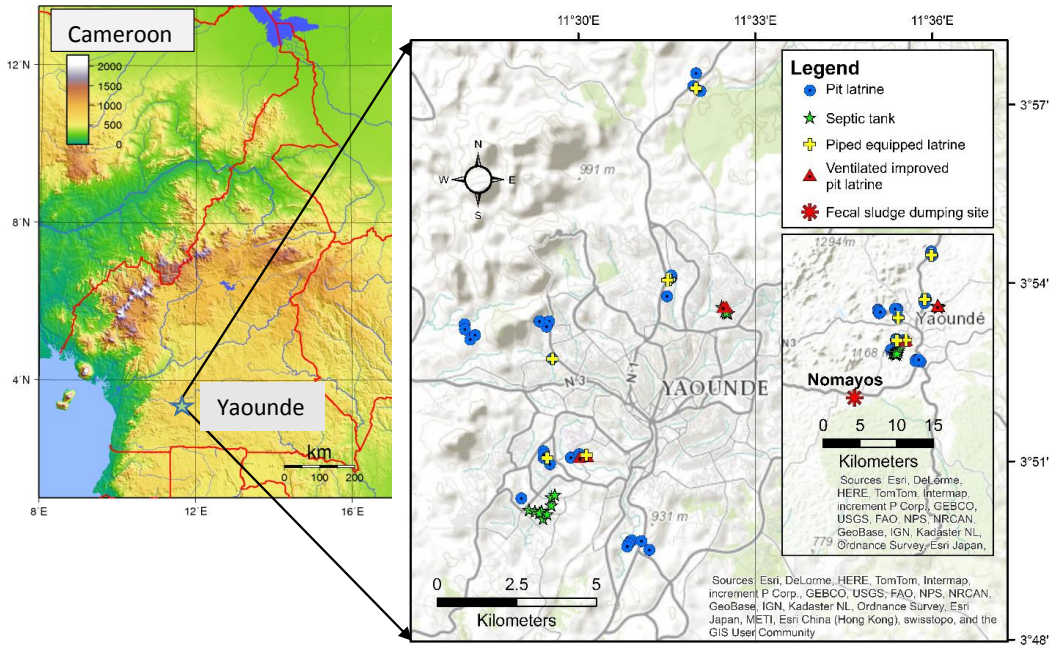
Figure 1: Map showing the location of Yaounde in Cameroon and the distribution of the different sanitation systems investigated in the study area as well as the discharge area of Nomayos.
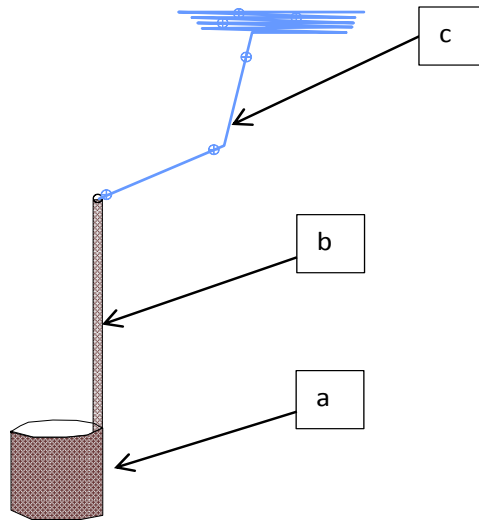

Figure 2: Faecal sludge sampling material needed to be adapted for this study: (a) the metal bar, (b) the metallic box, (c) the graduated rope.

Table 1: Summary of the descriptive statistic for the faecal sludge parameters analysed in this study.

Parameters	N	Mean	Sd	SE (Mean)	Minimum	Maximum	IQR	CV	Skewness
Temperature	53	27.82	1.66	0.23	22.80	30.30	1.90	0.05	-0.82
pH	53	7.25	0.71	0.09	5.90	8.94	1.04	0.09	0.41
EC [mS/cm]	53	5.46	3.91	0.53	0.22	17.38	4.58	0.72	1.20
Salinity [‰]	53	2.68	2.21	0.30	1.16	10.10	2.81	0.82	1.16
NH_4^+ [g/L]	53	0.57	0.55	0.07	0.005	2.48	0.76	0.97	1.45
TKN [g/L]	53	1.71	1.51	0.20	0.04	4.87	2.76	0.88	0.77
COD [g/L]	53	36.89	34.16	4.69	0.36	124.18	42.30	0.92	0.96
BOD_5 [g/L]	53	4.54	4.27	0.58	0.12	18.54	5.00	0.93	1.63
TS [%]	53	9.68	5.65	0.77	0.16	17.85	10.63	0.58	-49.00
TVS [%]	53	71.50	23.60	3.24	0.44	97.53	12.5	0.33	-1.85
TSS [g/L]	53	137.35	126.93	17.43	1.80	546.00	132.60	0.92	1.48
Water content [%]	53	58.40	7.64	1.04	42.12	76.5	9.47	0.13	0.22
HE (/L of FS)	53	2358	2308	320	33	8,000	2,783	0.98	1.07

Sd: standard deviation, SE (Mean): standard error of the mean, IQR: interquartile range, CV: coefficient of variation, n=number of samples analysed

RESULTS
Principal Component Analysis (PCA)

PCA was applied to the FS data to compare the compositional pattern between the analysed samples and to identify factors that influence each one. The number of components to keep was based on the Kaiser criterion, for which only the components with eigenvalues greater than 1 are retained. PCA of the data set brought out six principal components (PCs) with eigenvalues > 1 explaining about 78% of the total variance in the FS quality data set (Figure 3).

The first component accounting for 28.38% of the total variance was positively correlated (loading > 50) with COD, BOD_5, NH_4^+, TKN, TS and total helminth eggs (Table 2). The second principal component accounting for 18.54% of the total variance was positively correlated with electric conductivity and salinity. Analysis revealed that the first factor is generally correlated with the more variables than the second factor. The first factor indicated that the faecal sludge is characterized by their organic and parasitological fraction while the second factor is characterized by the physical fraction. Looking the third, fourth and sixth components (not showed in the Table 2 because of eigenvalues <1) accounted for the total variance of 13.31%, 9.66% and 8.13%.

Hierarchical Cluster Analysis (HCA)

The main result of the HCA performed on the 53 FS samples is shown in the dendrogram (Figure 4). This dendrogram classifies the samples into 9 clusters. Observation of the dendrogram revealed some indications of the level of similarity between the nine clusters (height). It showed 72.91% of variation within the samples of the same class and 27.09% between the different classes. It is important to mention that the graph was formatted in order to the better visualization of the different clusters, and therefore the maximal value of dissimilarity was set to 12%.

The characteristic of each cluster according to the 13 analysed FS parameters used in the HCA are described (Table 3). The Figure 4 indicates the presence of the many types of latrines into the same cluster stating the overlap in the quality of samples from different origin. Analyses of the variance between cluster showed significant variations for the clusters 4, 6 and 9 (Table 3). Furthermore, the samples from cluster 1 (C1) and cluster 3 (C3) are linked to the other clusters with an elevate distance between centroid classes; indicates that the samples are distinct from the other seven clusters (Figure 4). Among the seven other clusters, the cluster 2 has a high distance between centroid classes to the clusters C4, C5, C6, C7 and C9. The lower linkage distances (i.e. low

dissimilarities) were recorded between clusters C5 and C7; C4 and C9. In other words, the clusters C5 and C7, C4 and C9 have the greater similarity. It can be concluded that the clusters with lower distance between centroid classes are expected to content the FS of similar composition according to the physicochemical and parasitological parameters analysed. This information could allow concluding that the classification of FS in function of their origin (latrine types) could not constitute a supplement method for the assessment of the variability of FS as dissimilarities are observed in FS of the same origin.

The pH and the temperature of all clusters were not significantly different for the two first factors components of the factor analysis (Figure 5). The samples of cluster 1 (C1) and the cluster 3 (C3) were characterized by the highest concentration of EC, salinity, COD, NH_4^+, TKN, TS, TSS, total HE and water content contrary to the samples of the clusters 4 and 9.

Effects of latrine typology on the quality and variability of faecal sludge

The results obtained revealed the variability of the faecal sludge quality in samples. This variability was observed within the samples of the same origin and in the samples of different origin (Figure 6 and Figure 7). Assessment of the variability of physical parameters is showed in the Figure 6. The pH varied between the samples of the same groups without significant difference between the groups of latrines. The values were ranged from 6 to 9 units (Figure 6a). The distribution of pH range values in function of the type of latrine were 5.90 - 8.94, 6.76 - 7.21, 6.60 - 7.21 and 6,18 - 7.66 units respectively for traditional pit latrines, septic tanks, ventilated improved pit latrines and piped equipped latrines. The distribution of the electrical conductivity (EC) and salinity in function of latrine types have followed similar pattern. Samples from septic tanks (ST) were less concentrated than other samples from traditional pit latrines, ventilated improved pit latrines and piped equipped latrines (Figure 6b and 6d). The distribution of median values in function of latrines types were 4.40 mS/cm, 2.89 mS/cm, 4.35 mS/cm, 5.67 mS/cm and 2.40 ‰, 0.88 ‰, 2.37 ‰, 4.20 ‰ respectively

for traditional pit latrines, septic tanks, ventilated improved pit latrines and piped equipped latrines. Furthermore, the concentration of dry matters in samples was found to be inversely proportional to the water content. The distribution of median values recorded for dry matters were 5.29%, 9.68%, 12.79% and 14.70% respectively for septic tanks, piped equipped latrines, traditional pit latrines and ventilated improved pit latrines (Figure 6c). The distribution of organic contents represented by COD, BOD_5 and TKN (Figure 7) showed a heterogeneous distribution like other previous parameters. The samples were characterized by a very high COD content with the median values of 34 gO_2/L, 6.51 gO_2/L, 70.38 gO_2/L, and 31.81 gO_2/L respectively for traditional pit latrines, septic tanks, ventilated improved pit latrines and piped equipped latrines. In comparison to COD, the values of BOD_5 were very low with the median distributed as 3.20 gO_2/L, 1.53 gO_2/L, 6.50 gO_2/L, and 3.40 gO_2/L respectively (Figure 7a and 7c). Under the light of obtained results, the global median of the COD/BOD_5 mass ratio have varied between 1.15 and 70.08 in samples. Thus the COD/BOD_5 mass ratio greater than 3 was observed in the samples from traditional pit latrines (mean=7.54), ventilated improved pit latrines (mean =7.89) and piped equipped latrines (mean= 8.88). The septic tank was recorded to the lower COD/BOD_5 mass ratio with the value of 2.87.

The variation of the parameters TKN and NH_4^+ had followed similar patterns like other parameters described above. The median values recorded for TKN and NH_4^+ in function of latrine types were 1.73 g/L and 0.33 g/L, 0.25 g/L and 0.20 g/L, 2.81 g/L and 0.28 g/L, 0.84 g/L and 0.47 g/L respectively for the samples from traditional pit latrines, septic tanks, ventilated improved pit latrines and piped equipped latrines (Figure 7b and 7d).

Analysis of the relationship between both parameters taking into consideration the 53 samples and the type of latrines using the Pearson correlation test revealed that there are positive relationships between some organic parameters tested. The more significant correlations were those observed between organic parameters COD and BOD_5 TKN and nitrogen ammonia (NH_4^+) with respective

Pearson correlation score (R^2) of 0.618, 0.517, 0.782 and 0.338 (Figure 8). One other relevant outcome of the Pearson correlation matrix is the positive relationship observed between total helminth eggs and solid fraction of faecal sludge.

The prevalence distribution of helminth eggs (per litre of FS) in the different sanitation systems is shown in Figure 9. The septic tank and VIP latrine samples were less concentrated than traditional pit latrines and piped equipped latrine samples with respective mean values of total helminth eggs per litre of sludge of 1,150 and 1,960. The helminth egg species prevalence (mean values and standard errors) found in samples were *Ankylostoma duodenale* (14,797±2,825 eggs/L), *Ascaris lumbricoides* (16,852±1,905 eggs/L), *Enterorobius vermicularis* (12,202±2,015 eggs/L), *Fasciola hepatica* (14,969±2,529 eggs/L), *Hymenolepis nana* (10,171±1,838 eggs/L), *Schistosoma mansoni* (12,706±2,601 eggs/L), *Strongyloides stercoralis* (15,159±2,680 eggs/L), *Teania sp* (13,542 ±1,906 eggs/L), *Trichuris trichiura* (13,055±1,957 eggs/L).

The results of the Kruskal-Wallis statistical test indicated that many of physicochemical parameters varied between groups of latrines (Table 4). It is the case for the parameters salinity, TKN, DCO, TSS and Water content with respective p-values of 0.08, 0.03, 0.02, 0.01, 0.05 and 0.002. The non-significant variation of many other parameters between latrine groups was also recorded (p-values calculated > 0.05).

The classification of faecal sludge sampled in this study is shown on Table 5. The Table 5 indicated that FS sampled in this study can be classified into two types in accordance to the physico-chemical and parasitological parameters considered in the classification: the "high strength FS" with higher concentration and the "low strength FS" more less concentrate. However, there is an overlapping between qualities of samples from different latrines. The samples of the same origin are sometimes belonging to the different classes of FS. The mean feature of this classification is that FS samples from traditional pit latrines and ventilated improved pit latrines are more concentrated and they, for the majority belonging to the high strength FS. The samples from septic tanks were found to be more less concentrated and belonged to the low strength FS.
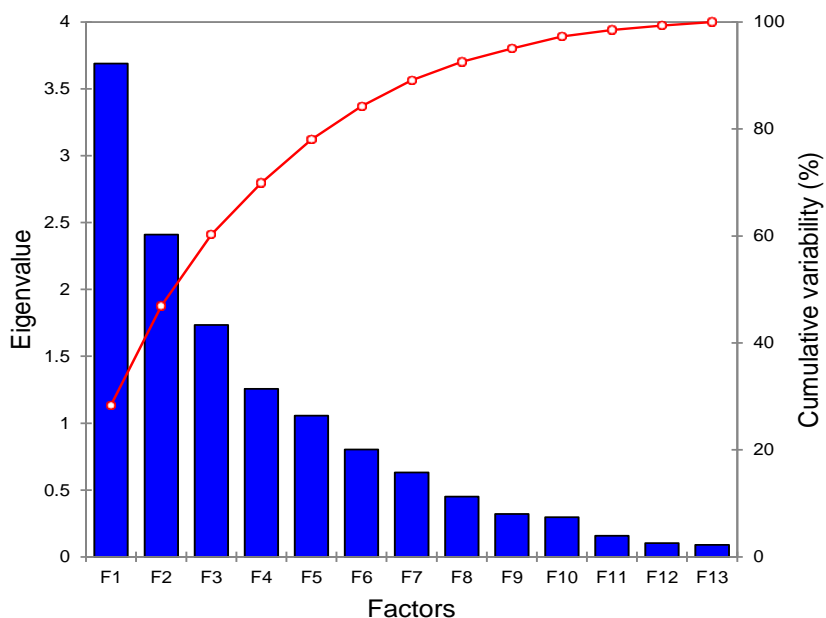

Figure 3: Scree plot showing the variation of eigenvalues and the cumulative variability of the total sample for the 13 faecal sludge parameters analysed.

Table 2: Factors loading for analysed parameters after varimax rotation.

Parameters	F1	F2
Temp [°C]	0.04	-0.80*
pH	-0.18	-0.65*
EC [mS/cm]	0.18	0.71*
Salinity [‰]	0.05	0.80*
COD [g/l]	0.79*	0.06
BOD_5 [g/l]	0.62*	-0.11
NH_4^+[g/l]	0.70*	0.15
TKN [g/l]	0.83*	-0.05
TVS [%]	-0.06	0.53*
TS [%]	0.73*	0.17
TSS [g/l]	0.52	0.26
Water content [%]	0.13	-0.25
HE [L of FS]	0.56*	0.07
Eigenvalue	3.69	2.41
Variance	28.38	18.54

* Absolute loading values > 0.50
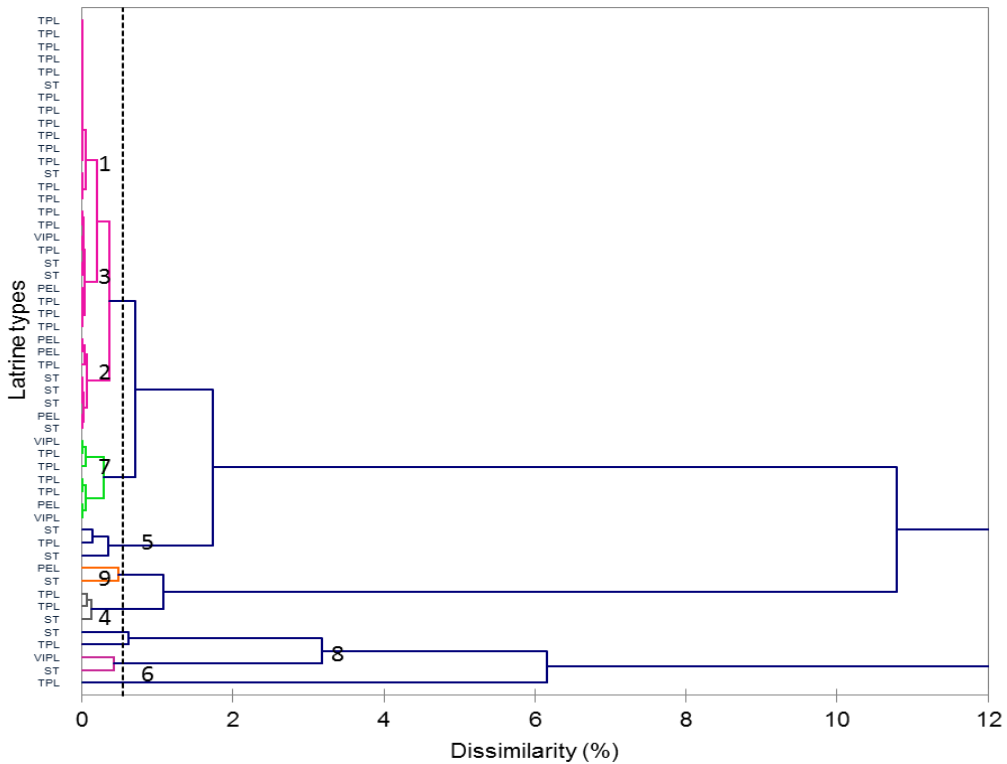
Figure 4: Dendrogram showing hierarchical classification of the 53 faecal sludge samples (in Arab number the identification and distribution of clusters): (PEL) piped equipped latrines, (ST) septic tank, (TPL) traditional pit latrine, (VIP) Ventilated improved pit latrine.

Table 3: Analysis of variance between clusters.

Cluster numbers	DF	Sum of squares	Mean squares	R²	F	Pr > F	F-ratio	Pr > F
Cluster 1	10	11.731	1.173	0.022	0.115	0.893		
Cluster 2	10	11.988	1.199	0.001	0.005	0.995		
Cluster 3	10	11.721	1.172	0.023	0.119	0.889		
Cluster 4	9	6.140	0.682	0.488	2.864	0.096	7.852	**0.021**
Cluster 5	10	11.823	1.182	0.015	0.075	0.929		
Cluster 6	9	6.918	0.769	0.424	2.204	0.157	5.592	**0.042**
Cluster 7	10	11.996	1.200	0.000	0.002	0.998		
Cluster 8	10	11.745	1.175	0.021	0.108	0.898		
Cluster 9	9	6.283	0.698	0.476	2.730	0.106	7.950	**0.020**

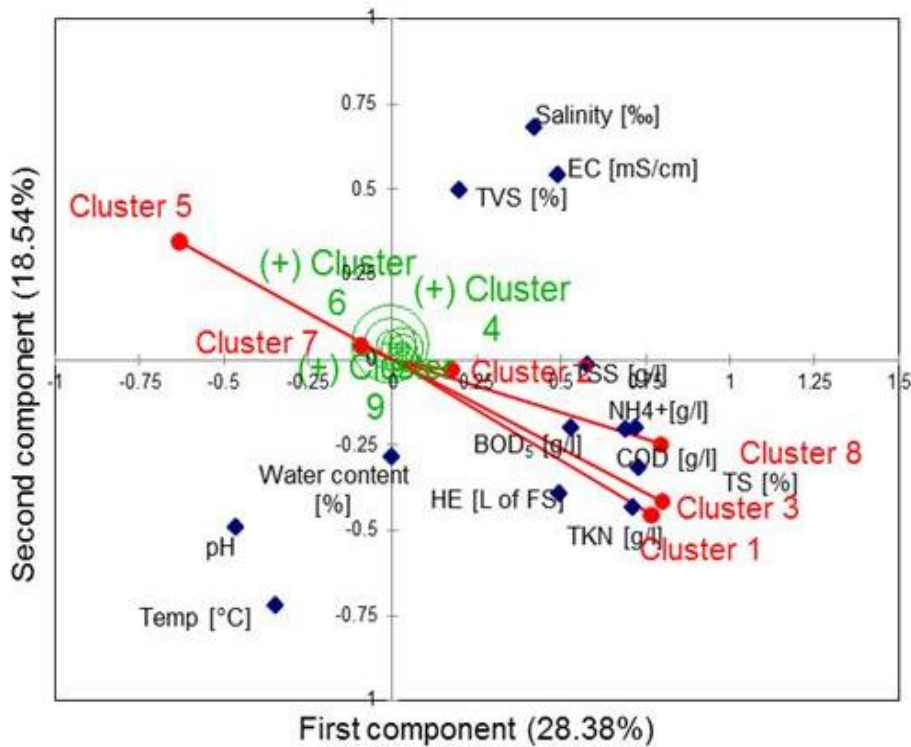DF: degrees of freedom; *Fisher* -statistic; Pr: probability value.

Figure 5: Preference map of the distribution of scores and loadings of the first two components of the principal component analysis showing the relationship between the physico-chemical parameters, the total helminth eggs content of the 53 samples as well as the 9 clusters of faecal sludge.
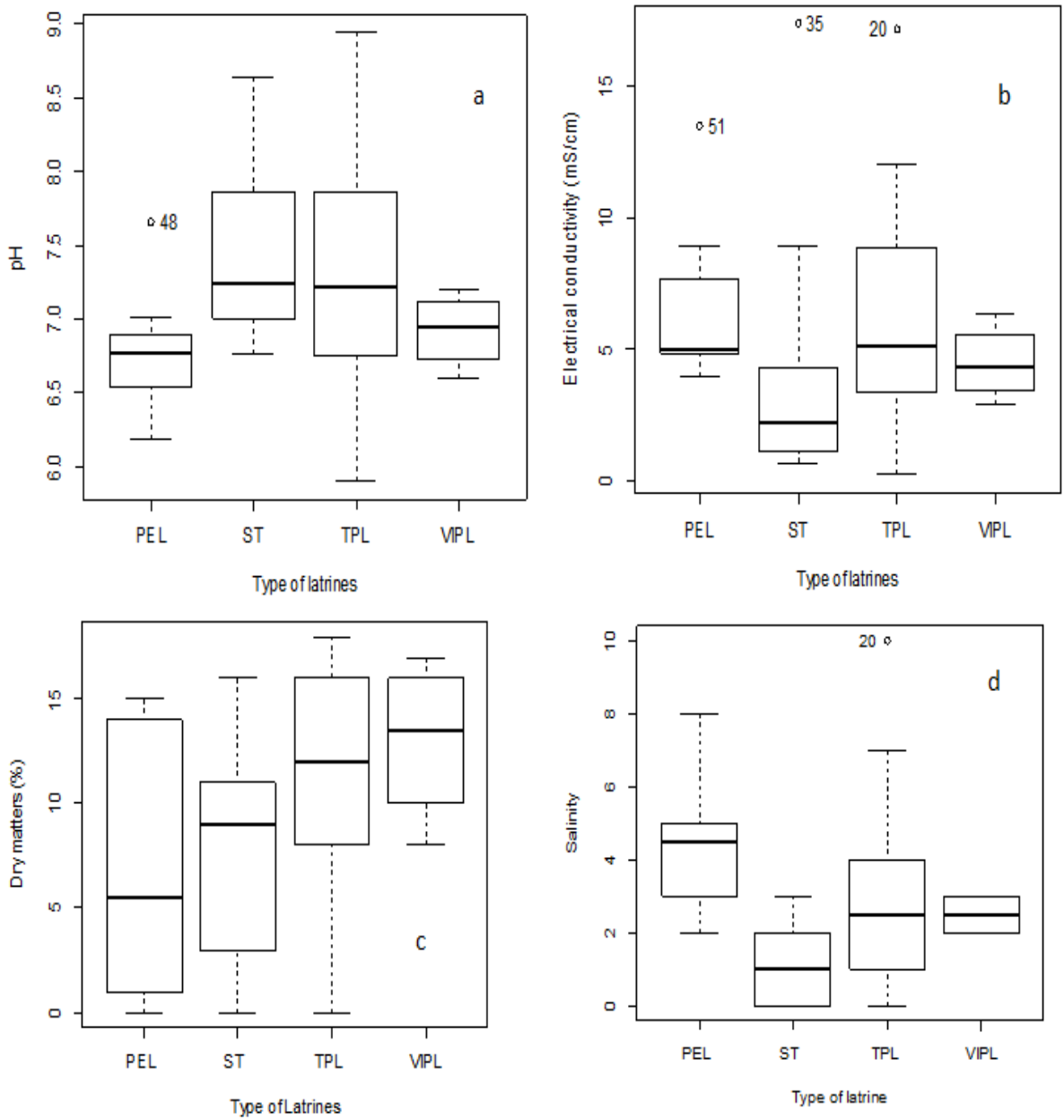
Figure 6: Box plot showing the variation of some physical parameters of faecal sludge in function of the type of latrines: (a) pH, (b) electrical conductivity, (c) dry matters, (d) salinity, (PEL) piped equipped latrines, (ST) septic tank, (TPL) traditional latrine, (VIPL) Ventilated improved pit latrine.
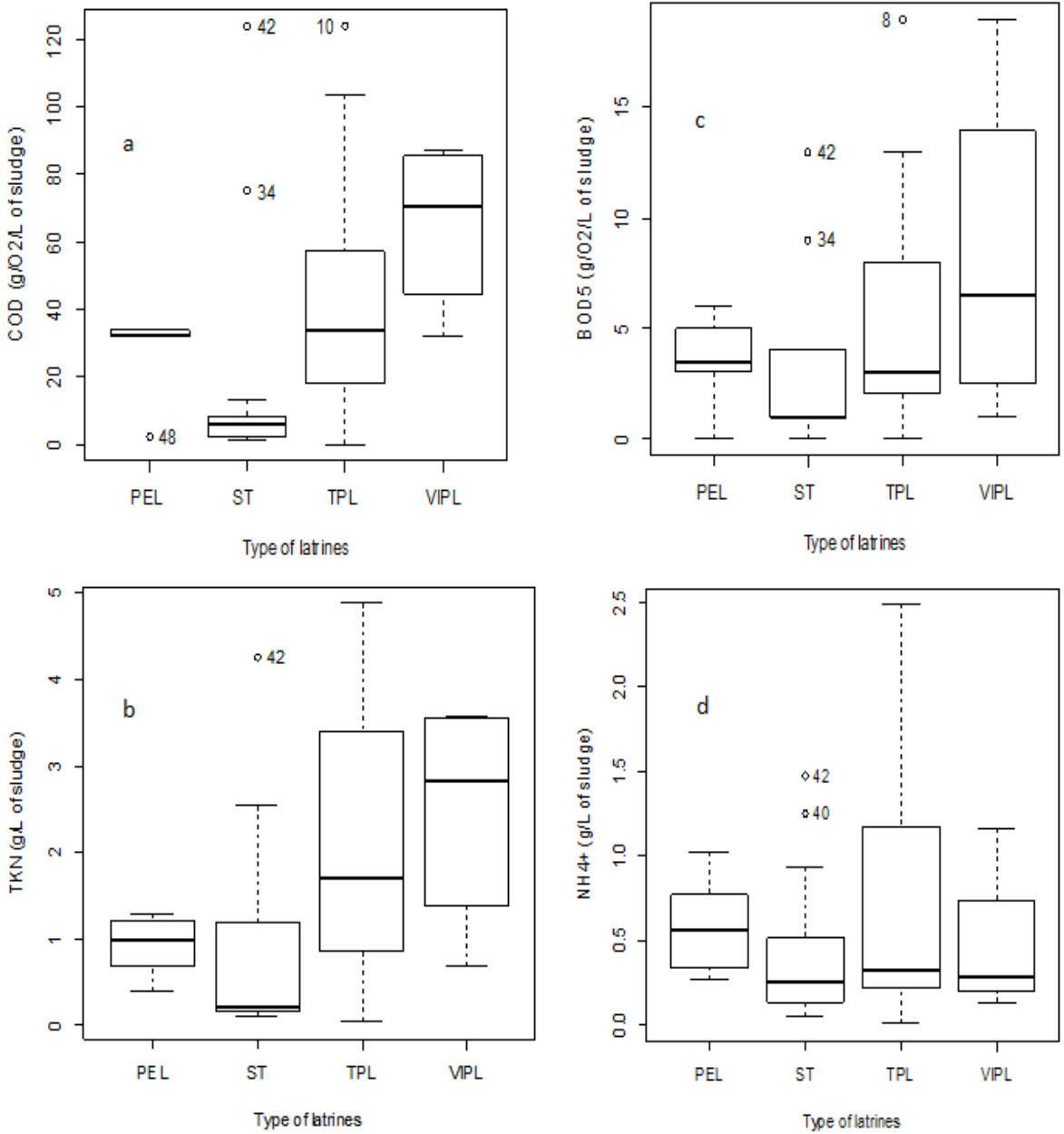
Figure 7: Box plot showing the variation of the organic and chemical contents of the FS samples: (a) COD, (b) TKN, (c) BOD, (d) NH_4, (PEL) piped equipped latrines, (ST) septic tank, (TPL) traditional latrine, (VIPL) Ventilated improved pit latrine (n=53).
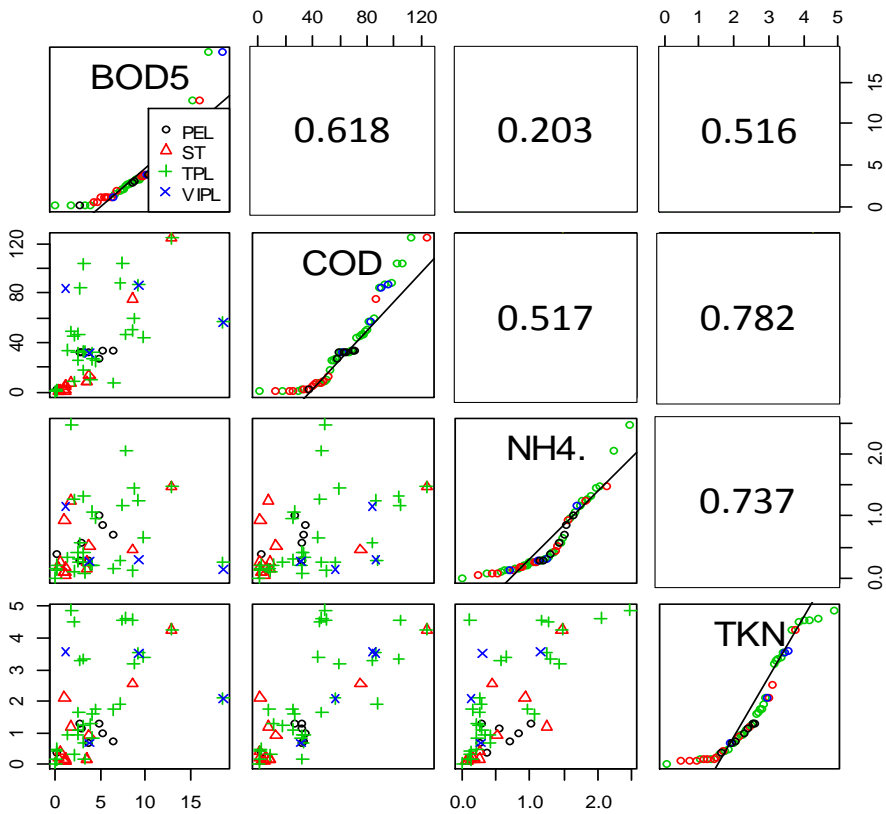
Figure 8: Cross correlation matrix (Pearson test, 95 % confidence level) showing relationship between BOD_5, COD, NH_4^+ and TKN of faecal sludge samples (PEL: piped equipped latrines; ST: septic tank; TPL: traditional pit latrine; VIP: ventilated improved pit latrine) (n=53).
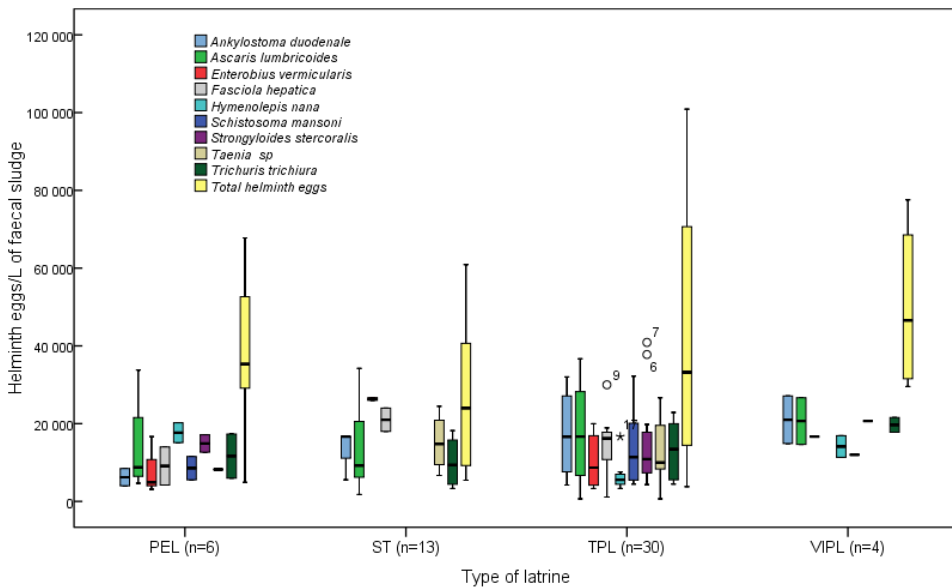

Figure 9: Box plot showing the variation of helminth egg diversities as well the total helminth eggs in samples (PEL: piped equipped latrines; ST: septic tank; TPL: traditional pit latrine; VIP: Ventilated improved pit latrine) (n=53).

Table 4: Summary results of the Kruskal-Wallis statistical test on the assessment of the variation of faecal sludge parameters between groups of latrines.

Parameters	X^2	df	p-value
Temperature	4.09	3	0.26
pH	6.1	3	0.11
EC	6.55	3	0.08
Salinity	8.94	3	0.03*
NH_4^+	1.92	3	0.59
TKN	9.37	3	0.02*
DCO	10.6	3	0.01*
DBO_5	4.18	3	0.24
TS	5.163	3	0.16
TVS	4.8	3	0.19
TSS	7.11	3	0.05*
Water content	14.9	3	0.002*
Total HE (Eggs/L of FS)	5.92	3	0.12

X^2= value of chi-squared obtained for the one-way analysis of variance test; df= degrees of freedom.

Table 5: Classification of faecal sludge sampled in this study according to Heinss et al. (1998)*.

Parameters	Samples per type of latrines (Mean)				low strength FS*	High strength FS*
	PEL (n=6)	ST (n=13)	TPL (n=30)	VIPL (n=4)		
pH	6.69	7.24	7.18	6.95		
EC [µS/cm]	5.67	2.89	4.83	4.35		
Salinity [‰]	4.20	0.88	2.40	2.37		
COD [g/L]	31.81	6.51	34	70.38	<10	20.-50
BOD_5 [g/L]	3.40	1.53	3.20	6.50	7.6	0.84 - 2.6
COD/DBO_5					5:1-10:1	2:1-5:1
NH_4+ [g/L]	0.47	0.21	0.33	0.28	1.00	2.-5
TKN [g/L]	0.84	0.22	1.73	2.81		
TS (as % of TVS)	5.45	8.54	12.47	13.40	<3%	≥3%
TSS [g/L]	46.49	79.87	124.50	394.70	≈7	≥30
Total HE [L of FS]	33,968	25,641	41,4590	50,055	≈4,000	20-60,000

(PEL) Piped equipped latrine; (ST) Septic tank; (TPL) Traditional pit latrine; (VIP) Ventilated improved pit latrine; (HE) Helminth eggs; (n) number of samples.

DISCUSSION

Principal Component Analysis (PCA)

This study revealed that the components of the PCA account for 78% of variability in the FS quality data set. It revealed the variability in concentration of the FS samples collected in the context of Yaounde city. The variability of FS is discussed in the literature (Foxon et al., 2012; Strande et al., 2014). This variability may be due to the type of on-site sanitations in use at household level, householder feeding habits as well as the storage duration, the temperature, the intrusion of groundwater, the mixture with solid wastes and the tank emptying technologies. Variability was also observed in the field of water resource management. Indeed, observations carried out by Liu et al. (2003) using factors analysis for the interpretation of 28 groundwater samples, mentioned that 77.8% of the total groundwater quality variation for the two first factors model with eigenvalues >1. In addition, Cloutier et al. (2008) recorded 78.3% of variability for the first five components of the PCA when applying multivariate statistical analysis of geochemical data for the groundwater evolution in a sedimentary rock aquifer system.

Hierarchical Cluster Analysis (HCA)

Hierarchical Cluster Analysis showed 72.91% of variation within the samples of the same class and 27.09% between the different classes. The presence of the latrine diversities into the same cluster revealed the overlapping in the quality of FS samples from different on-site sanitation technologies. Analyses of the variance between clusters showed significant variations between the clusters. The result confirmed that the quality of FS have changed within the study area regardless of the source of the samples. Given the diversity of the treatment systems (Ingallinella et al., 2002), each class or group of FS could be adapted for an appropriate treatment technology. The classification of samples in water research studies using the HCA was demonstrated by several authors. Indeed, Cloutier et al. (2008) analysed geochemical data (14 parameters) as

indicative of hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system in Canada. The authors classified 153 groundwater samples into 7 clusters grouping the samples of similar characteristics according to the distribution of hydro geochemical parameters and they integrated the HCA for the regional division of the aquifer system. In addition, Wang et al. (2013) by assessing the surface water quality data (15 parameters) collected from 6 sites of the Songhua River Harbin region (China) classified the 6 sites into 3 clusters and concluded that the whole six monitoring sites on the river were grouped into three statistically distinctive groups.

Effects of latrine typology on the quality and variability of faecal sludge

The sanitation technologies in use at the household level have affected the quality and variability of FS. The variation of FS pH was also recorded by several authors (Ingallinella et al., 2002; Cofie et al., 2006; Kengne et al., 2008; Soh Kengne et al., 2014; Letah Nzouebet et al., 2016). This variability of the pH between samples could depend on the variation of environmental conditions (temperature, humidity, availability of oxygen), the type and the rate of biological process occurring in FS storage technologies. Indeed, Strande et al. (2014) describe the pH as an essential parameter to understand FS chemistry process such as acid-base chemistry and biological process. The low values of EC and salinity recorded for septic tanks could be explained by the high mineralization of FS occurring in septic systems (Montangero and Belevi, 2007). The septic tank systems could be more favourable to the better stabilization of FS during storage. Dry matter of samples has widely varied with higher values for samples from traditional pit latrines and VIP latrines. The heterogeneity in the distribution of dry matter in samples can be explained by the on-site sanitations use pathways because septic tanks and piped equipped latrines operate as wet systems and require time to time water for flushing urine and faecal matters into the pit thereby contributing to the

dilution of the concentration of dry matters recorded in these systems. Indeed, the variability of FS moisture content (inversely proportional to dry matters) was observed by several authors (Radford and Sugden, 2014; Bakare et al., 2012). Indeed, the experiments carried out by Radford and Sugden (2014) working on the characterization and fluidization of synthetic pit latrines sludge revealed a variation of FS water content.

The low degradability (COD/DBO$_5$ ratio >3) of the samples belonging to traditional pit latrines, ventilated improved pit latrines, piped equipped latrines could have resulted to the poor activities of microorganisms involved in the mineralization process of the faecal matter into the pit of sanitation technologies. The low microbial degradability could also be explained by the latrines usages. Indeed, Still et al. (2012) mentioned the negative effects of disinfectants containing sodium hypochlorite added to the latrine pits to reduce odours. The authors pointed out the effects of disinfectants in the inhibition of the metabolism of heterotrophic bacteria which were active in sludge degradation. Furthermore, many of the latrines investigated in this study, particularly traditional pit latrines were containing household solid waste refuses (anal cleansing materials, plastic wastes, menstrual hygiene material …) that could negatively affect the biodegradability of FS by reducing the surface contact between microorganisms and faeces. In the same way of observation, Bakare et al. (2012) working on the variation of VIP latrine sludge contents also found the quality of faecal sludge to be depended on household habits combined to local environmental conditions. Variations were also observed in the Nitrogen forms. Indeed, Fidjeland et al. (2013) demonstrated that ammonia (NH$_3$) in faecal sludge is produced by deamination of organic nitrogen and hydrolysis of urea (CO(NH$_2$)$_2$) in urine by urease.

The diversity of helminth eggs found in FS samples corroborates the study of Jimenez et al. (2000). The explanation of this finding could be that the septic tanks and VIP latrines are more suitable for helminth eggs

which dies-off. Furthermore, the inactivation processes of helminth eggs could be the result of environmental factors such as unavailable oxygen, production of ammonia (NH$_3$), effect of temperature combined to the storage duration. Indeed, Pecson and Nelson (2005) and Pecson et al. (2007) demonstrated the positive effects of temperature and the ammonia concentration in the inactivation of *Ascaris* eggs in sewage sludge. Additionally, the variation of the total helminth eggs in the study area could also be explained by the difference in the prevalence of helminthic diseases in the study area because the health status of latrines users could thereby affect helminth egg frequencies content in faecal sludge. The variation of the total helminth eggs in faecal sludge samples were found by several authors (Jimenez, 2007; Seck et al., 2015). For Ingallinella et al. (2002), many areas of Africa, Asia, and Latin America are known to present the higher prevalence of nematode infections, specifically the helminth infections.

Conclusion

The purpose of the study was to investigate the variation in the physico-chemical and parasitic composition of the on-site sanitation faecal sludge by using two multivariate statistical methods, the hierarchical clustering analysis (HCA) and the principal component analysis (PCA). The HCA helped to group the 53 FS samples into 09 clusters including samples of similar characteristics. The PCA of the whole data set denoted six PCs and explained for about 78% of the data variability. The significant effect of the on-site sanitation type on quality and variability of FS were recorded. The samples from septic tanks were less concentrated than other samples originating from traditional pit latrines, ventilated improved pit latrines and piped equipped latrines expressing the importance of types of sanitation technologies in the variability of faecal sludge quality data. PCA and HCA could constitute effective methods of analysing, interpreting and representing data concerning the faecal sludge quality. However, the use of multivariate

statistical methods in the description of data from the field of the FS characterization needed to be conducted worldwide in order to more document and integrate these methods in the literature of the FS characterisation. This could help in the practicing of faecal sludge management in the orientation and choice of the FS treatment systems based on quality data of the classes or clusters after analysis.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

WALN designed of the study, wrote the protocol, elaborated the faecal sludge sampling methodology used on the field and laboratory analysis of samples, analysis of data and wrote the first draft of the manuscript. ESK, GVDW, AJNN and PN contributed to faecal sludge sampling in the field and performed laboratory analysis. CH contributed to analysis of data and scientific orientation of the manuscript. AR reviewed the manuscript and its scientific orientation. IMKN reviewed the sampling methodology, reviewed the manuscript as well as its scientific orientation. All authors read and approved the final manuscript.

ACKNOWLEDGEMENTS

REFERENCES

Acar EF, Sun L. 2013. A generalized Kruskal–Wallis test incorporating group uncertainty with application to genetic association studies. *Biometrics,* **69**: 427–435.

Ambassa N. 2005. Contribution à la caractérisation de la production des boues de vidange dans la ville de Yaoundé. Master thesis, Faculty of Science, University of Yaounde I, Cameroon. 45 p.

Anderberg MR. 2014. *Cluster Analysis for Applications: Probability and Mathematical Statistics: A Series of Monographs and Textbooks.* Academic Press.

Anh TH, Koottatep T, Polprasert C. 2018. Business model analysis for faecal sludge collection and transport services in Thailand and Vietnam. *Journal of Water, Sanitation and Hygiene for Development,* **8**(3): 556-567. DOI: 10.2166/washdev2018.026.

APHA/AWWA/WEF. 2005. *Standard Methods for the Examination of Water and Wastewater* (21st Edn). American Public Health Association, and Water and Environment Federation Publication: Washinton D.C., USA.

Bakare BF, Foxon KM, Brouckaert CJ, Buckley CA. 2012. Variation in VIP latrine sludge contents. *Water SA,* **38**: 479–486.

Bassan M, Tchonda T, Yiougo L, Zoellig H, Mahamane I, Mbéguéré M, Strande L. 2013. Characterization of faecal sludge during dry and rainy seasons in Ouagadougou, Burkina Faso. In: Proceedings of the 36th WEDC International Conference, Nakuru, Kenya. pp. 1–5.

Bastos VK, Cutolo SA, Doria M, Razzolini MTP. 2013. Detection and quantification of viable *Ascaris sp.* and other helminth eggs in sewage sludge. *Int. J. Environ. Health Res.*, **23**: 352–362.

Berteigne B. 2012. Quantification and characterization of faecal sludge in the cities of Yaounde and Douala, Cameroon. Master thesis, National

School of Engineers of Water and Environment, Strasbourg, France, 112p.

BUCREP. 2012. Results of the general population census in Cameroon. Technical Report of BUCREP, http://www.bucrep.cm, p 36.

Chatfied C, Collins AJ. 2013. *Introduction to Multivariate Analysis.* Springer.

Cloutier V, Lefebvre R, Therrien R, Savard MM. 2008. Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. *J. Hydrol.,* **353**: 294–313.

Cofie OO, Agbottah S, Strauss M, Esseku H, Montangero A, Awuah E, Kone D. 2006. Solid–liquid separation of faecal sludge using drying beds in Ghana: Implications for nutrient recycling in urban agriculture. *Water Res.,* **40**: 75–82.

Fidjeland J, Magri ME, Jönsson H, Albihn A, Vinner B. 2013. The potential for self-sanitisation of faecal sludge by intrinsic ammonia. *Water Res.,* **47**: 6014–6023.

Foxon K, Brouckaert C, Wood K. 2012. How fast do pits fill up? Empirical evidence and mathematical models. FSM2, Durban, South Africa.

Green PE. 2014. *Mathematical Tools for Applied Multivariate Analysis*. Academic Press.

Heinss U, Larmie SA, Strauss M. 1998. Solids Separation and Pond Systems. TheTreatment Faecal Sludges Trop. Lessons Learnt Recomm. Prelim. Des.

Hothorn T, Everitt BS. 2014. *A Handbook of Statistical Analyses Using R*. CRC Press.

Ingallinella AM, Sanguinetti G, Koottatep T, Montangero A, Strauss M, Jimenez B, Spinosa L, Odegaard H, Lee DJ. 2002. The challenge of faecal sludge management in urban areas- strategies, regulations and treatment options. *Water Sci. Technol.,* **46**: 285–294.

Jimenez B. 2007. Helminths (Worms) eggs control in wastewater and sludge. In: International Symposium on New Directions in Urban Water Management. pp. 12–14.

Kaiser HF. 1958. The varimax criterion for analytic rotation in factor analysis. *Psychometrika,* **23**: 187–200.

Kengne IM, Dodane P-H, Akoa A, Koné D. 2009. Vertical-flow constructed wetlands as sustainable sanitation approach for faecal sludge dewatering in developing countries. *Desalination,* **248**: 291–297.

Kengne IM, Akoa A, Soh EK, Tsama V, Ngoutane MM, Dodane PH, Koné,D. 2008. Effects of faecal sludge application on growth characteristics and chemical composition of *Echinochloa pyramidalis* (Lam.) Hitch. and Chase and *Cyperus papyrus* L. *Ecol. Eng.,* **34**: 233–242.

Kengne IM, Dodane PH, Akoa A, Koné D. 2009. Vertical-flow constructed wetlands as sustainable sanitation approach for faecal sludge dewatering in developing countries. *Desalination,* **248**: 291–297.

Koné D, Strauss M, 2004. Low-cost options for treating faecal sludge (FS) in developing countries–Challenges and performance. In: 9th International IWA Specialist Group Conference on Wetlands Systems for Water Pollution Control and to the 6th International IWA Specialist Group Conference on Waste Stabilisation Ponds, Avignon, France.

Kone M, Service E, Ouattara Y, Ouattara P, Bonou L, Joly P. 2016. Caractérisation des boues de vidange dépotées sur les lits de séchage de zagtouli (Ouagadougou). *Int. J. Biol. Chem. Sci.,* **10**(6): 2781-2795.

Letah Nzouebet WA, Kengne Noumsi IM, Rechenburg A. 2016. Prevalence and diversity of intestinal helminth eggs in pit latrine sludge of a tropical urban area. *J. Water Sanit. Hyg. Dev.,* **6**: 622–630.

Lienou G, Mahe G, Paturel JE, Servat E, Sighomnou D, Ekodeck GE, Dezetter A, Dieulin C. 2008. Evolution des régimes hydrologiques en région équatoriale camerounaise: un impact de la variabilité climatique en Afrique équatoriale? *Hydrol. Sci. J.,* **53**: 789–801.

Liu CW, Lin KH, Kuo YM. 2003. Application of factor analysis in the assessment of groundwater quality in a blackfoot

disease area in Taiwan. *Sci. Total Environ.,* **313**: 77–89.

Liu X, Laverty DM, Best RJ, Li K, Morrow DJ, McLoone S. 2015. Principal Component Analysis of Wide-Area Phasor Measurements for Islanding Detection—A Geometric View. *Power Deliv. IEEE Trans.,* **30**: 976–985.

Montangero A, Belevi H. 2007. Assessing nutrient flows in septic tanks by eliciting expert judgement: A promising method in the context of developing countries. *Water Res.,* **41**: 1052–1064.

Niwagaba CB, Mbéguéré M, Strande L. 2014. Faecal Sludge Quantification, Characterisation and Treatment Objectives. *Faecal Sludge Manag. Syst. Approach Implement. Oper.,* **19**.

Parrot L, Sotamenou J, Dia BK. 2009. Municipal solid waste management in Africa: Strategies and livelihoods in Yaoundé, Cameroon. *Waste Manag.,* **29**: 986–995.

Patras A, Brunton NP, Downey G, Rawson A, Warriner K, Gernigon G. 2011. Application of principal component and hierarchical cluster analysis to classify fruits and vegetables commonly consumed in Ireland based on in vitro antioxidant activity. *J. Food Compos. Anal.,* **24**: 250–256.

Pecson BM, Barrios JA, Jiménez BE, Nelson KL. 2007. The effects of temperature, pH, and ammonia concentration on the inactivation of Ascaris eggs in sewage sludge. *Water Res.,* **41**: 2893–2902.

Pecson BM, Nelson KL. 2005. Inactivation of Ascaris suum eggs by ammonia. *Environ. Sci. Technol.,* **39**: 7909–7914.

Radford JT, Sugden S. 2014. Measurement of faecal sludge in-situ shear strength and density. *Water SA,* **40**: 183–188.

Seck A, Gold M, Niang S, Mbéguéré M, Diop C, Strande L. 2015. Faecal sludge drying beds: increasing drying rates for fuel resource recovery in Sub-Saharan Africa. *J. Water Sanit. Hyg. Dev.,* **5**: 72–80.

Shrestha S, Kazama F. 2007. Assessment of surface water quality using multivariate statistical techniques: A case study of the Fuji river basin, Japan. *Environ. Model. Softw.,* **22**: 464–475.

Singh EJ, Gupta A, Singh NR. 2013. Groundwater quality in Imphal West district, Manipur, India, with multivariate statistical analysis of data. *Environ. Sci. Pollut. Res.,* **20**: 2421–2434.

Soh Kengne E, Kengne IM, Nguetsop VF, Foubi SI, Akoa A, Strande L. 2014. Algal diversity and distribution in waste stabilization ponds treating faecal sludge leachate from drying vegetated beds. *Int. J. Biol. Chem. Sci.,* **8**(3): 946-955.

Still D, Louton B, Bakare B, Taylor C, Foxon KM, Lorentz SA. 2012. Investigating the Potential of Deep Row Entrenchment of Pit Latrine and Waste Water Sludge for Forestry and Land Rehabilitation Purposes: Report to the Water Research Commission. Water Research Commission.

Strande L, Ronteltap M, Brdjanovic D. 2014. *Faecal Sludge Management: Systems Approach for Implementation and Operation*. IWA Publishing.

Strauss M, Larmie SA, Heinss U. 1997. Treatment of sludges from on-site sanitation—low-cost options. *Water Sci. Technol.,* **35**: 129–136.

Strauss M, Montangero A. 2002. Feacal Sludge Management Review of Practices, Problems and Initiatives. Dubendorf EAWAGSANDEC.

Tilley E, Ulrich L, Lüthi C, Reymond P, Zurbrügg C. 2014. *Compendium of Sanitation Systems and Technologies*. Eawag.

Wang Y, Wang P, Bai Y, Tian Z, Li J, Shao X, Mustavich LF, Li BL. 2013. Assessment of surface water quality via multivariate statistical techniques: A case study of the Songhua River Harbin region, China. *J. Hydro-Environ. Res.,* **7**: 30–40.