Statistical Evaluation of Influence of Hydrochemical Parameters on Public Water Supply in Ibadan, Nigeria

O. S. Oladeji[1] B. S. Aribisala[2]
1.Department of Civil Engineering, Ladoke Akintola University of Technology, Ogbomoso, Nigeria
2.Department of Computer Science, Lagos State University, Lagos, Nigeria
Corresponding author: osoladeji@lautech.edu.ng; olayinka_oladeji@yahoo.co.uk

Abstract
The recent trends in the global socio-economic and technological developments have exerted undue pressure on water quality and supply. A robust model that predicts water supply from hydrochemical parameters could be useful in water resources management. Most of the existing models used bivariate analysis or simple multivariate analysis that treats each variable separately, but combined parameters could be better at capturing many related water properties and hence have stronger predictive value. In this work, exploratory factor analysis was used to investigate the relationships between fourteen measured hydrochemical parameters and the volume of water pumped for public water supply from Asejire waterworks in Ibadan. Representative monthly samples of raw water were collected and analyzed for their physical and chemical properties. Principal Component Analysis (PCA) was used to derive five novel latent variables which represent combined properties of water. The results showed that, with the exception of colour, no other single hydrochemical parameter was significantly associated with the amount of water pumped for public supply. Interestingly, one of the novel latent variables 'Opaque Index' was significantly associated with water supply. Additionally, though colour as a single parameter had a predictive value, but it explained only 6.7 % of the variation in the amount of water pumped, whereas the Opaque Index explained 25.5 % of the variation in the amount of water pumped. This therefore suggests that the Opaque Index latent variable has a stronger predictive power than any of the individual measured parameters and hence could a useful metric for effective water resources management.
Keywords: Hydrochemical, Latent variable, Prediction, Structural equation modelling

1 Introduction

Recent trends in the global socio-economic and technological developments coupled with the need for ecological conservation and effects of climate change, as well as the increasing water consumption rate relative to actual population growth have subjected the natural water resources regime to immerse pressure, thus limiting the availability of the resource (Diakité *et. al.,* 2009). In Nigeria, the annual urbanization growth rate for major cities is between 10 – 15 % (Yusuf, 2007) and approximately 63 million (about 40%) of the population are not within the reach of improved source of safe drinking water (UNICEF/WHO, 2012).

Ibadan is a Metropolitan city in the south western part of Nigeria (Figure 1), with population growth of approximately 300 % over a 33-year period (1963 – 1996). This growth exerted undue pressure on the public utilities including water quality and supply. Therefore, an efficient public water supply will require integrated management of both the water and land regimes.

The commencement of the operation of public water supply systems in Nigeria predates the Country's independence. However, despite immense investment of capital expenditure, significant development has been held back by poor water supply infrastructure, lack of technical capabilities, inefficient landuse planning and regulatory framework. A prototype for public water supply in Nigeria includes abstraction of surface water, treatment and distribution. The choice of a suitable treatment process, its design and operations are largely determined by the quality of the source water, which in turn determines the required infrastructure and corresponding cost. Source water laden with high contaminant concentration will potentially require enhanced chemical consumption or new infrastructural support, thus increasing the overall cost of treatment without corresponding increase in the amount of portable water produced.

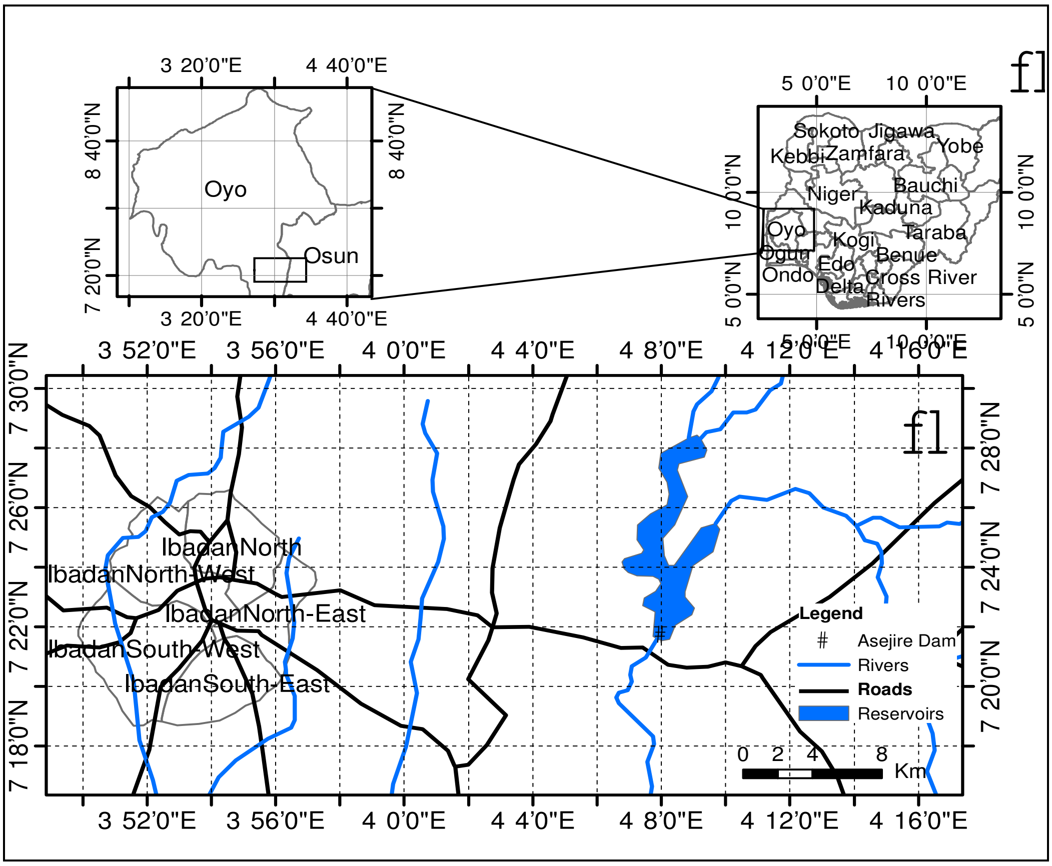

Figure 1 Location of the study area

Kanade and Gaikwad (2011) applied computational analysis on hydrochemical data obtained from Nashik and Nipahd taluka in India, to establish strong positive correlation between the individual parameters and subsequently inferred strong evidence of anthropogenic activities. Bajpayee *et. al*. (2012) applied Pearson correlation matrix and multiple linear regression analysis to hydrochemical data of ground water samples collected from north-eastern part of Bankura District, West Bengal, India to determine interactions between the individual parameters of ground water and also established strong positive relationships with the Total Dissolved Solids (TDS). Several other authors (Fan *et. al*., 2010; Wang *et. al*., 2013; Adetunji and Odetokun, 2011; Pinto and Maheshwari, 2011) have applied computational analysis to investigate temporal and spatial variations in water quality and identify sources of water pollution. These authors used bivariate analysis which suggests that the individual hydrochemical parameters are linearly related and hence using a single parameter may not be strongly reflective of the true characteristics of water sample in a particular location.

The aim of this study was to use exploratory statistical approach to investigate if the use of multiple parameters could offer a better representation of water properties than the use of each single parameter. The relationships between certain hydrochemical parameters and the amount of portable water produced over a 10-year period from Asejire waterworks were tested. Five novel latent variables were derived from the measured hydrochemical parameters. In addition, associations between the latent variables and the volume of portable water were also tested. The capability of each of the latent variable in predicting corresponding volume of public water supply was also investigated. The authors hypothesized that at least one of the latent variables will have a predictive power and will also have a stronger predictive power than each of the measured single water property.

2.0 Methodology

Representative monthly samples of raw water from Asejire Dam (Figure 1) were collected over a 10-year period (2000 – 2009), and analyzed for 14 hydrochemical parameters namely colour, turbidity, dissolve oxygen, pH, alkalinity, hardness (total, Mg and Ca). Others include odour, chloride, silica, chlorine residual and chlorine demand, as well as flocculation. The samples were collected using pre-cleaned 2-litre plastic containers (with lids), appropriately labeled and delivered to the laboratory for analyses. In addition, the volume of the pumped water for public supply over the same period of research was acquired from the Water Corporation of Oyo State, Nigeria.

Field observations and measurements were carried out for the turbidity, dissolved oxygen, pH and alkalinity, using field meters. The true color was determined by first removing all suspended substances through filtration and then measuring the color of the water sample by visual comparison with a series of specific color scale using spectrophotometer. The turbidity was measured using Nephelometer. The total alkalinity (as $CaCO_3$ in mg/L) was also determined using Hanna portable pH meter and H_2SO_4 standard titration method. The total hardness (mg/L as $CaCO_3$) was determined using EDTA titrant and standard laboratory titrimetric equipment. The chloride concentration was also measured using silver nitrate solution titration method.

The statistical analyses were performed using SPSS version 21.0 (IBM SPSS Inc. New York, USA), with all statistical tests being two-tailed, and P values < 0.05 being considered statistically significant. Bivariate associations using Pearson correlation were assessed among the hydrochemical parameters as well as between each of the parameters and the corresponding water volume pumped for public supply. Multivariate associations were investigated using factor analysis. Firstly, five latent variables were derived from the hydrochemical parameters using Principal Component Analysis (PCA). Only variables that were positively correlated with each other were used for the derivation of the latent variables. Secondly, associations between each of the derived latent variables and the corresponding water pumped for public supply were evaluated. Latent variables are commonly employed in statistical modelling when the parameter of interest is inadequately measured by a single variable. Latent variables are estimated based on the shared common variance between a set of measured indicator variables. This allows exclusion of measurement errors when estimating the contribution that each measured variable makes to the latent variable (Stephenson *et al*., 2006). This approach allowed the assessment of relationships between the water volume pumped for public supply and the corresponding values of groups of hydrochemical parameters of the raw water samples.

3.0 Results and discussions

The descriptive statistics of the measured hydrochemical parameters as well as the volume of the pumped water supply is presented in Table 1. The bivariate analysis (Table 2) showed that colour was significantly positively associated with turbidity and dissolved oxygen, but significantly negatively associated with the total alkalinity and total hardness. Higher turbidity was significantly associated with lower total hardness while higher pH values were significantly associated with higher concentration of Chloride. The results also indicate that higher alkalinity values were significantly associated with higher values of total and calcium hardness, and correspondingly to lower values of chlorine demand. Also, higher chlorine demand was significantly associated with lower amount of Silica and lower degree of calcium hardness. Higher values of total hardness were significantly associated with higher degree of calcium and magnesium hardness, as well as to the higher values of chloride, chlorine demand and correspondingly to lower values of silica. The bivariate analysis shows that only colour was significantly negatively associated with the amount of water pump, and no other hydrochemical parameters show significant association with the volume of water pumped for public water supply.

Table 1: Descriptive statistics of the hydrochemical parameters and pumped water supply

Parameter	Mean ± std
Colour	90.87 ±61.88
Turbidity	27.09 ±31.08
Dissolved oxygen	11.94 ±16.19
pH	7.21 ±0.25
Total alkalinity	55.03 ±12.46
Total Hardness	52.53 ±13.41
Calcium Hardness	35.55 ±12.2
Magnesium Hardness	16.04 ±10.97
Chloride	20.65 ±8.11
Silica	10.21 ±3.75
Chlorine Demand	10.81 ±18.4
Water Pumped	1045805.73 ±361641.76

In the multivariate analysis, several combinations of hydrochemical parameters were investigated using exploratory factor analysis. PCA was used to derive some latent variables based on those parameters that are linearly related. Only the latent variables with excellent factor loadings on the measured parameters are retained and reported in this work. Colour, Turbidity and Dissolved Oxygen were all positively correlated and therefore a latent variable termed 'Opaque Index' was derived from these measured parameters. The Opaque Index accounted for 84.1% of variance in colour, 68.2% of variance in turbidity and 60.1% of variance in dissolved oxygen. Similar approach was used to derive four additional latent variables, having established positive significant associations between some groups of measured hydrochemical parameters. The latent variable 'Tart Index' was derived from pH, chloride and total hardness and this variable accounted for 53.2% of variance in pH, 81.5% of variance in chloride and 64.2% of variance in total hardness. The latent variable 'TCA' was derived from total hardness, calcium hardness and total alkalinity. The TCA accounted for 78.1% of variance in total hardness, 82.3% of variance in calcium hardness and 75.9% of variance in total alkalinity.

Also, the latent variable 'Hardness Index' was derived from total hardness, calcium hardness and magnesium hardness and accounted for 91.2% of variance in total hardness, 68.5% of variance in calcium hardness and 55.0% of variance in magnesium hardness. Finally on latent variable derivation, 'TCC' was derived from total hardness, calcium hardness and chloride, and this accounted for 84.5% of variance in total hardness, 78.0% of variance in calcium hardness and 54.8% of variance in chloride. These factor loadings which ranged from 54.8% to 91.2% showed that the latent variables gave excellent fit to the measured variables.

The multivariate analysis (Table 3) showed that higher scores of the Opaque Index was significantly associated with lower amount of water pump but no significant association was found between water pump and other latent variables. The results of bivariable analysis show that with the exception of colour, all the measured parameters were not significantly associated with the amount of water pumped for public supply. However, the derived Opaque Index latent variable shows significant association with the amount of water pumped. The result showed further that the Opaque Index explained 25.5% of the variations in the amount of water pumped for public supply. This therefore supports the authors' hypothesis that at least one latent variable derived from the measured hydrochemical parameters is able to predict the amount of water pumped for public water supply.

Table 2: Pearson bivariate Correlations (p values) within measures of water properties and between measures of water properties and the amount of water pump

Parameter	Colour	Turbidity	Dissolved oxygen	pH	Total alkalinity	Total Hardness	Calcium Hardness	Magnesium Hardness	Chloride	Silica	Chlorine Demand
Turbidity	**0.43 (<0.001)**										
Dissolved Oxygen	**0.33 (0.001)**	0.08 (0.512)									
pH	0.05 (0.598)	0.09 (0.405)	-0.16 (0.115)								
Total alkalinity	**-0.27 (0.006)**	-0.16 (0.151)	0.12 (0.259)	0.03 (0.748)							
Total Hardness	**-0.32 (0.001)**	**-0.27 (0.014)**	-0.08 (0.422)	0.02 (0.873)	**0.37 (<0.001)**						
Calcium hardness	-0.14 (0.162)	-0.07 (0.53)	0.19 (0.072)	-0.15 (0.124)	**0.45 (<0.001)**	**0.48 (<0.0001)**					
Magnesium hardness	-0.11 (0.264)	-0.15 (0.193)	0.1 (0.337)	-0.12 (0.232)	0.14 (0.172)	**0.4 (<0.0001)**	-0.05 (0.633)				
Chloride	0 (0.997)	-0.03 (0.799)	0 (0.975)	**0.24 (0.025)**	0 (0.985)	**0.3 (0.004)**	0.16 (0.145)	-0.04 (0.74)			
Silica	0.03 (0.779)	0.19 (0.094)	0.1 (0.338)	0 (0.97)	0.17 (0.096)	**-0.24 (0.015)**	0.08 (0.43)	-0.03 (0.765)	-0.12 (0.258)		
Chlorine Demand	-0.07 (0.507)	-0.18 (0.117)	-0.08 (0.414)	0.16 (0.113)	**-0.23 (0.022)**	**0.24 (0.015)**	**-0.38 (<0.001)**	-0.13 (0.203)	0.06 (0.592)	**-0.26 (0.008)**	
Water Pumped	**-0.26 (0.023)**	-0.17 (0.172)	-0.13 (0.261)	0.03 (0.764)	-0.01 (0.937)	0.11 (0.319)	-0.14 (0.202)	0.08 (0.498)	-0.13 (0.279)	0.04 (0.726)	0.17 (0.13)

Note. Boldface represents associations that are significant at p<0.05

The colour parameter explained only 6.7% of the variation in the amount of water pumped whereas the Opaque Index latent variable explained 25.5% of the variations in the amount of water pumped, which therefore suggests that the Opaque Index latent variable has a stronger predictive power than any of the individual measured parameters as hypothesized.

Table 3:Association between latent variables (derived from measured hydrochemical parameters) and the volume of pumped water using regression analysis.

Latent Variable	Beta (p)	r squared
Opaque Index	-0.25 (0.02)	0.255
Tart Index	0.00 (0.969)	0.004
TCA	-0.02 (0.858)	0.02
Hardness Index	0.03 (0.791)	0.03
TCC	-0.05 (0.673)	0.047

Note. Values are standardized beta values (p) predicting pumped water from each latent variable. 'Opaque Index' = latent variable for colour, turbidity and dissolved Oxygen. 'Tart Index' = latent variable for pH, chloride and total hardness. 'TCA' = latent variable derived from total hardness, calcium hardness and total alkalinity. 'Hardness Index' = latent variable derived from total hardness, calcium hardness and magnesium hardness. 'TCC'= latent variable derived from total hardness, calcium hardness and chloride.

4 Conclusions

This work assesses several combinations of hydrochemical parameters using exploratory factor analysis and derived five novel latent variables, and those with excellent factor loadings on the measured parameters were reported. With the exception of colour, all the measured hydrochemical parameters were not significantly associated with the amount of water pumped for public supply. However, the colour parameter explained only 6.7% of the variation in the amount of water pumped whereas the novel Opaque Index latent variable explained 25.5% of the variations in the amount of water pumped, which therefore suggests that the Opaque Index latent variable has a stronger predictive power than any of the individual measured parameters. These findings suggest, as hypothesized, that at least one of the latent variables has a predictive power and also that the latent variable has a stronger predictive power than each of the measured single water property. Hence, this work concludes that the use of the latent variable which captures different but related water properties could improve water supply prediction and management.

References

Adetunji, V. O. and Odetokun, I. A., 2011. Groundwater contamination in Agbowo community, Ibadan Nigeria: Impact of septic tanks distances to wells. Malaysian Journal of Microbiology, 7(3); 159-166.

Bajpayee, S., Das, R., Ruj, B., Adhikari, K., and Chatterjee, P.K., 2012. Assessment by multivariate statistical analysis of ground water geochemical data of Bankura, India. International Journal of Environmental Sciences. 3(2); 870 – 880.

Diakité, D., Semenov, A. and Thomas, A., 2009. A proposal for social pricing of water supply in Côte d'Ivoire. Journal of Development Economics 88; 258–268

Fan, X., Cui, B., Zhao, H., Zhang, Z., and himing, Zhang, H., 2010. Assessment of river water quality in Pearl River Delta using multivariate statistical techniques. Procedia Environmental Sciences 2; 1220–1234

Kanade S. B. and Gaikwad V. B., 2011. A Multivariate Statistical Analysis of Bore Well Chemistry Data - Nashik and Niphad Taluka of Maharashtra, India. Universal Journal of Environmental Research and Technology. 1 (2); 193-202.

Pinto, U., and Maheshwari, B.L., 2011. River health assessment in peri-urban landscapes: An application of multivariate analysis to identify the key variables. Water Research. 45; 3915 – 3924.

Stephenson M, Holbert R, Zimmerman R., 2006. On the use of structural equation modeling in health communication research. Health Commun; 20:159-167.

UNICEF and World Health Organization (UNICEF/WHO), 2012. Progress on Drinking Water 2012 UPDATE. Joint Monitoring Programme for Water Supply and Sanitation. ISBN: 978-924-1503297; 61p.

Wang, Y., Wang, P., Bai, Y., Tian, Z., Li, J., Shao, X., Mustavich, L. F. and Li, B., 2013. Assessment of surface water quality via multivariate statistical techniques: A case study of the Songhua River Harbin region, China. Journal of Hydro-environment Research 7; 30 – 40.

Yusuf, K. A., 2007. Evaluation of groundwater quality characterisitics in Lagos-City. Journal of Applied Sciences 7(13), 1780-1784.

The IISTE is a pioneer in the Open-Access hosting service and academic event management. The aim of the firm is Accelerating Global Knowledge Sharing.

More information about the firm can be found on the homepage:
http://www.iiste.org

CALL FOR JOURNAL PAPERS

There are more than 30 peer-reviewed academic journals hosted under the hosting platform.

Prospective authors of journals can find the submission instruction on the following page: http://www.iiste.org/journals/ All the journals articles are available online to the readers all over the world without financial, legal, or technical barriers other than those inseparable from gaining access to the internet itself. Paper version of the journals is also available upon request of readers and authors.

MORE RESOURCES

Book publication information: http://www.iiste.org/book/

Academic conference: http://www.iiste.org/conference/upcoming-conferences-call-for-paper/

IISTE Knowledge Sharing Partners

EBSCO, Index Copernicus, Ulrich's Periodicals Directory, JournalTOCS, PKP Open Archives Harvester, Bielefeld Academic Search Engine, Elektronische Zeitschriftenbibliothek EZB, Open J-Gate, OCLC WorldCat, Universe Digtial Library , NewJour, Google Scholar