Regression Based Modelling of Relationship Between Lead in Water and Lead Bioaccumulation in Fish

O.K. Adeyemo, O.A. Banwo and E.A. Adeleye

Department of Veterinary Public Health and Preventive Medicine, University of Ibadan, Nigeria
Department of Geography, University of Ibadan, Nigeria

Abstract: Aquatic ecosystems around the world, lake, estuaries and coastal areas are increasingly impacted by anthropogenic pollutants through different sources such as agricultural, industrial and urban discharges, atmospheric deposition and terrestrial drainage. This study tested the hypotheses that Level of lead in water is significantly related to level of lead accumulated in fish tissue. GIS analysis and multiple regression analysis were used to build a relationship between the lead concentration in fish and the lead concentration in water. The results from this study revealed lead levels higher (0.05-0.07 mg kg⁻¹) than WHO’s recommended maximum lead levels of 0.01 mg kg⁻¹ in fishes. The results (r² = 0.612, P = 0.066) of the analyses imply that there is a strongly positive correlation between the lead in the water and the quantity of lead accumulated in fish tissue. The regression model shows that lead concentration in fish is accounted for by 61% of lead concentration in water. The regression model is therefore statistically significant at p>0.05. To fill the gap when no fish samples are taken and to maintain a periodic real-time lead accumulation forecast in the sampled location, the regression model generated is therefore an efficient tool. Moreover, by integrating GIS with the simulated results, the visualization technique is helpful for rapid understanding of water quality conditions. Regression model developed in one city could also be readily applied to other locations and other years when recalibrated against data from a few local monitoring sites, with comparable results [1].

Key words: Modelling • Lead • Bioaccumulation • Fish • pollution

INTRODUCTION

Water resources are at the heart of sustainable development in many regions of the world [2]. Surface waters such as rivers, streams and ponds are extremely important for human existence because water of sufficient quantity and quality is an essential resource for agriculture, industry and tourism and also for everyday life in cities and villages. Human populations depend on water for drinking, agriculture, maintenance of livestock and fishing. Therefore, whether viewed as resource or commodity, water is the basis of our agricultural, municipal, industrial, environmental and aesthetic well-being, but it has been alternately taken for granted, abused and exploited [3].

Pollution has become an increasing environmental, economic and social problem for many communities around the world, extreme numbers of lakes, streams, rivers, ponds and watersheds have all been affected by water pollution from human impacts on the environment [4]. Sources of pollution, effects of pollution and predicting pollution are all topics that scientists and policy-makers alike have concerned themselves with [5]. In metropolitan cities like Ibadan, pollutants from different sources end up directly or indirectly in the water bodies. A considerable amount of pollution is from domestic and industrial effluents [3]. Pollutants have also been linked to run off water from mechanic workshops and car wash shops which are usually located near rivers and streams. One of the most ubiquitous contaminant in these runoff waters is lead. A GIS is a computer database and graphics system which can perform spatial analytical functions with geographically referenced data [6]. GIS can map land use and land cover changes over time, which can then give information about the rate, character and site of land use change. This becomes useful when determining non-point sources of pollution into water bodies [7].

Simple regression is used to examine the relationship between one dependent and one independent variable. After performing an analysis, the regression statistics can be used to predict the dependent variable when the
independent variable is known [8]. Regression goes beyond correlation by adding prediction capabilities.

One of the major challenges in epidemiological research is to devise appropriate metrics and methods for exposure assessment. In the context of non-point surface water pollution, this is particularly problematic because of continuing uncertainty about the source of pollutant, the likelihood of important interactive and cumulative effects from different pollutants, high levels of both spatial and temporal variability in pollutant concentrations and a dearth of monitoring data. Against this background, models that can estimate at unsampled locations are clearly needed. This paper presents an example of how Geographic Information System (GIS) techniques can be used to develop such models for urban-scale analysis, on the basis of readily available data. This study assessed the lead level in randomly selected rivers and muscles of fish caught from such rivers. From the results, a predictive model will be developed, which will be able to determine the corresponding lead level in fish muscle based on the knowledge of the lead level in sampled surface water.

**MATERIAL AND METHODS**

**Study area:** This scope of this study includes the major rivers in Ibadan city. The river systems in Ibadan studied include the main rivers which are the Ona River, the Ogunpa River and the Ogbere River. The other rivers are tributaries of these main ones. Ibadan is the capital of Oyo state of Nigeria. It has an elevation of about 210m above sea level and it is located on latitude 7° 26’N and longitude 3° 5’E. The area occupied by the metropolitan area of the city is drained by the Ogunpa River and Oga River. Ogunpa River is a 3rd order stream. Ogunpa River drains the Eastern side while Oga River drains the Western side of Ibadan.

**Water samples:** Water samples were taken from Asejire dam, Ogunpa River, Eleyele River, Oga River; the samples from Oga were taken from three different locations (Oga a, Oga b and Oga c) within the metropolis. Mid-stream water samples were collected from the rivers at each sample point and stored in pre-cleaned 500ml plastic containers before being taken to the laboratory for analysis.

**Fish samples:** Six samples of fish were taken from the six rivers. The species of fish sampled is Tilapia (*Oreochromis niloticus*) the fish samples were collected live from the rivers by the fishermen. There were kept in the river water in well ventilated plastic containers and transported to the laboratory.

**Procedure for dry ashing and determination of lead in the fish sample:** Fish samples were weighed in its fresh state. The weight was recorded as $W_1$.

The sample was then oven dried and the weight determined again; this is the dry weight recorded as $W_2$. Then, the moisture content was determined as follows: $W_1 - W_2$.

The dried sample was afterwards grinded into pieces. 0.2 gram of the grinded samples was placed in a crucible of known weight ($W_3$). This crucible with grinded sample was placed in the furnace at 550°C until the sample was burnt into ashes. After a shing, the crucible with its content was removed and placed inside a desiccator and allowed to cool. The crucible with the ashed sample was then weighed ($W_{cs}$).

The Ash weight can then be determined as thus

$$\text{% Ash content} = \frac{W_{cs} - W_c}{W_{cs}}$$

**Determination of lead content of samples:** The lead content of the samples was determined by Atomic Absorption Spectrophotometer.

**Statistical analysis:** The results of the lead level in the sampled rivers and fish muscles were computed. Logistic regression, correlation tests and ANOVA were used for statistical analysis of the data using Microsoft Excel software.

**Regression models:** A Regression model was developed from the data. The model was used to build the relationship between the lead concentration in fish and the lead concentration in water. The significance of the model was then determined.

**Geographic Information System (GIS):** A digital map of Ibadan that covers the eleven (11) local government area was derived using ArcView 3.1. The data of lead in water and fish muscles were imputed in GIS (ArcView 3.1) software and maps were generated.

**RESULTS AND DISCUSSION**

A major barrier to accurate short-term forecasts is the lack of an efficient system for water quality monitoring. Traditional water quality sampling is time-consuming, expensive and can only be done for small areas. In terms of performance, regression models have a relatively good record. Where they have been validated against
Fig. 1: Lead level in water and fish muscle from the different rivers sampled

Fig. 2: Map of Ibadan, the study area
Fig. 3: Map of Ibadan region showing sample points to be yellow after values has been imputed
MAP OF IBADAN
Sample Points
- Lead in Fish Tissue
- LGA Boundaries
- Ibadan Rivers
- Lakes
- Landuse Types
- Light Forest
- Savannah Woodland
- Open Grassland
- Low Shrub
- Forest Reserve
- Builtup Area
- Airport Runway

Fig. 4: Map showing a GIS model bar graph indicating the level of lead contamination of sampled surface water.

Fig. 5: Map showing a GIS model bar graph indicating level of lead bioaccumulation in fish tissue.
Table 1: Pearson correlation relationship between lead in water and lead in fish

<table>
<thead>
<tr>
<th></th>
<th>Lead concentration in fish</th>
<th>Lead concentration in river</th>
</tr>
</thead>
<tbody>
<tr>
<td>Lead concentration in fish</td>
<td>Pearson correlation</td>
<td>1</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>N</td>
<td>6</td>
<td>6</td>
</tr>
<tr>
<td>Lead concentration in river</td>
<td>Pearson correlation</td>
<td>0.782</td>
</tr>
<tr>
<td>Sig. (2-tailed)</td>
<td>.</td>
<td>.</td>
</tr>
<tr>
<td>N</td>
<td>6</td>
<td>6</td>
</tr>
</tbody>
</table>

Table 2: Model summary of the variables in the equation and regression

<table>
<thead>
<tr>
<th>Model</th>
<th>R</th>
<th>R square</th>
<th>Adjusted R square</th>
<th>Std. error of the estimate</th>
<th>Change statistics</th>
<th>df1</th>
<th>df2</th>
<th>Sig. F change</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.782a</td>
<td>0.612</td>
<td>0.514</td>
<td>0.005582</td>
<td>R square change</td>
<td>0.612</td>
<td>6.297</td>
<td>1</td>
</tr>
</tbody>
</table>

Table 3: Coefficient model of relationship between lead in water and lead in fish

<table>
<thead>
<tr>
<th>Model</th>
<th>Unstandardized coefficients B</th>
<th>Unstandardized coefficients std. Error</th>
<th>Standardized coefficient beta</th>
<th>t</th>
<th>Sig.</th>
</tr>
</thead>
<tbody>
<tr>
<td>(Constant) lead Concentration in rivers</td>
<td>3.399</td>
<td>0.011</td>
<td></td>
<td>3.167</td>
<td>0.036</td>
</tr>
<tr>
<td></td>
<td>4.130</td>
<td>0.016</td>
<td>.782</td>
<td>2.506</td>
<td>0.066</td>
</tr>
</tbody>
</table>

monitored data, Coefficients of determination ($r^2$) typically in the range of 0.45-0.7 and standard errors of <20% of the mean [9, 10, 11, 12, 13] have been reported; comparable to more sophisticated dispersion modeling (Briggs, 2005). In the present study, Lead level in the sampled surface water ranged between 0.5-0.90 mg L$^{-1}$ (mean = 0.65 mg L$^{-1}$), while lead level in muscles of fish sampled from the surface water ranged between 0.05-0.07 mg kg$^{-1}$ (mean = 0.06 mg kg$^{-1}$) (Figure 1). This result revealed that lead level is higher than WHO’s recommended maximum lead levels of 0.01 mg kg$^{-1}$ in fishes. The relationship between Lead concentration in fish tissue and Lead concentration in water assessed using Pearson correlation is presented in Table 1.

A single land use regression model developed in one city could readily be applied to other locations and other years when recalibrated against data from a few local monitoring sites, with comparable results [1]. Owing to the inter-correlations between many of the variables used in land use regression modeling, substitution of one measure with another, for the sake of comparability, will rarely alter model performance to any noticeable degree. The significance level (0.066) of Pearson correlation in the present study, shows that the relationship is positive and strong, but it is not significant at $p<0.05$. The regression model shows that lead concentration in fish is accounted for by 61% of lead concentration in water (Table 2). Clearly, regression methods can also be enhanced in several ways. One important improvement would be to recognize (and separately model) the different contributions from local and long-range (and primary and secondary) sources.

From the results of the coefficient analysis (Table 3), a model was generated thus:

$$Y = 0.034 + 0.0413X$$

Where $Y$ represents values of lead concentration in fish muscle and $X$ represents values of lead concentration in water.

Thus, if the value of lead level in water is known, the corresponding values in fish can be determined from this model. Likewise, the Lead amount that will accumulate in fish muscle for a unit increase in Lead Level of water can be determined as well. Thus, the model is adequate ($r = 0.782$, $r^2 = 0.612$, $P = 0.066$). The regression model is therefore statistically significant at $p<0.05$. Display sequential images of water quality predictions provide decision-makers easily understandable information. Correctly locating where the problems are occurring is the necessary first step in identifying how to correct the problem. GIS is often the best tool for this type of problem solving [14-16]. The map generated from this study provides solid basis for decision-makers to estimate the consequences of lead pollution event in all river section.
REFERENCES


