Groundwater Pollution Vulnerability Assessment Using A New GIS-Based DRASTIC Method

S.M. Hamza, A. Ahsan, M. Imteaz, T.A. Mohammad and A.H. Ghazali

1Department of Civil Engineering, University Putra Malaysia, UPM 43400, Serdang, Selangor, Malaysia
2Department of Civil Engineering, Hussaini Adamu Federal Polytechnic, PMB 5004 Kazaure, Jigawa state Nigeria
3Faculty of Engineering and Industrial Sciences, Swinburne University of Technology, Hawthorn, Melbourne, VIC 3122, Australia

Abstract: Protection of groundwater from pollution is generally a global political decision. These decisions are essentially supported by DRASTIC analysis-being the most popular method of evaluating the intrinsic pollution vulnerability of groundwater. However, because of the spatial variability around the world, the model has been exposed to various transformations. Especially the weights assigned to the parameters constituting the model, for their enormous effect towards achieving final vulnerability result. Popular among the DRASTIC weights transformation technique are the single parameter sensitivity analysis (SPSA) and analytic hierarchy process (AHP). While these methods are proven effective in some studies, they are contrarily ineffective in others. In this study a new approach was developed using Pearson’s product moving correlation (PPMC). The weight assigned to a parameter is directly proportional its level of input on attaining nitrate contamination. Hence it is likely to be most suitable for virtually all areas as the groundwater vulnerability map is generated based on the pollution pattern of the area under study. Seven models were compared to determine the best option for the study area. The nitrate validation result indicated that DRASTIC was 0.73, FRASTIC 0.65, DRASIC-SPSA 0.75, FRASTIC-SPSA 0.75, DRASTIC-AHP 0.74, FRASTIC-AHP 0.63 and FRASTIC-PPMC which is developed in this study was 0.77. As far as the study area used is concerned the new developed DRASTIC-PPMC had proven to be the most effective groundwater pollution vulnerability assessment approach.

Key words: Pearson’s product moving correlation • Sensitivity analysis • GIS • AHP • Kano

INTRODUCTION

Protection of ground water from pollution is commonly a global political decision and the decisions are essentially aided by DRASTIC analysis [1, 2, 3]. Global embracement of DRASTIC model has made it the most popular method of evaluating the intrinsic pollution vulnerability of groundwater to date [4-12]. Because it also provides satisfying results in vast areas with a multifaceted geological structure [13]. DRASTIC model was introduced by the American water well association for application by the environmental protection agency (EPA) for groundwater pollution assessment. DRASTIC is an abbreviation for seven (7) parameters adopted in the model, namely: D (depth of water), R (net recharge), A (aquifer media), S (soil media), T (topography), I (impact of the vadose zone) and C (hydraulic conductivity of the aquifer). The establishment of the parameters, ratings and weights were based on Delphi technique [14].

In an attempt to assess groundwater vulnerability around the world, DRASTIC model has undergoes various changes of its parameters weights and rating [3, 15, 17]. The flexible nature of the model allows for either inclusion or reduction from the principal parameters to suit a particular area [28, 29]. As part of the models accomplishment study by [30], it was reported that 50% of close to 1 million km² studied using DRASTIC falls within medium to high vulnerability.
Even though Delphi was employed in defining the weights assigned to parameters, the regional peculiarity necessitated adjustments. The weight assigned to a parameter plays a significant role in attaining vulnerability indices [31, 32]. This sprung the introduction of sensitivity analysis (SA). Single parameter sensitivity analysis (SPSA) was developed by [33]. SA is to date the most the acknowledged technique [1, 11, 32, 34-40]. SPSA is primarily employed to aid researchers in assessing the influence of each parameter towards attaining the level of vulnerability based on its Delphi assigned weight [41, 31]. SA is central to achieving correct interpretation of the resulting vulnerability [31, 32, 11, 42].

Another method employed by researchers for the model transformation is a technique for multi criteria decision making (MCDM) known as the analytic hierarchy process (AHP). It entails the employment of pair-wise comparison matrices (PCMs) for the study of various criteria. AHP is used to develop ratio scales from both discrete and continuous paired comparisons. It was developed by [43]. The pairs to be compared in the technique may be founded based on consensus or actual measurement [44]. A triumph of the AHP in decision making process is evident in a study by [12, 45-52].

Map validation is central for effective land use management [53]. The validation. For groundwater vulnerability maps validation, nitrate is usually used because it is absent in groundwater under natural condition and its presence indicates pollution [29]. The validity of maps is realized through comparison of various approaches with an applied contaminated situation [41], as apparent in a study by [17, 3, 54, 55, 46, 47, 56].

While these methods (SPSA and AHP) are proven to be effective in some studies, they are contrarily ineffective in others based on the achieved coefficient. This reason has prompted researchers in a quest for a better technique as evident in a study by [46] where frequency ratio technique was used.

This study presents an original pioneer approach for DRASTIC rates optimization. While other studies employed the nitrate only for model validation, this study, however, generates the parameter weights based on their relationship with the nitrate concentration. Pearson’s product moving correlation (PPMC) was employed. In this method, the weight assigned to a parameter is directly proportional its level of input towards attaining nitrate contamination. Hence it is likely to be most suitable for virtually all areas; as the groundwater vulnerability map is generated based on the pollution pattern of the area under study.

**MATERIALS AND METHODS**

**Study Area:** Kano is one of the 36 states of Nigeria located in the Sudan Savannah between latitude 10° 23’ 40” and 12° 34’ 24” North, 7° 41’ 15” and 9° 21’ 21” East with a total area cover of 20, 131 km² shown in Figure 1. Its climate is seasonally arid. Rain falls between May and October with peak in August and the mean annual rainfall is between 635 to 1000 mm.

**Data Collection:** The data used for the assessment of groundwater vulnerability is shown in Table 1. The parameter maps for the DRASTIC and F were produced within the GIS environment using spatial analysis.

**Methodology:** Figure 2 indicates the general methodology adopted in the development of the new groundwater vulnerability approach. The steps are itemized as follows:

- The standard DRASTIC methodology by [14] was employed: Since the hierarchy of the ratings and weights were established based on Delphi technique, the ratings were maintained in both models.
- Nitrate data is interpolated using Kriging interpolation technique.
- Points are created. The points are spread evenly so as to have a better representation of the entire study area.
- The parameters in the DRASTIC were also interpolated. While interpolating the standard DRASTIC ratings are maintained.
- Values corresponding to the location of the points created were extracted for all the parameter maps including the nitrate map.
- All the extracted values were exported to Microsoft Excel for evaluation.
- The correlation between the nitrate and each of the DRASTIC parameters was obtained.
- The correlation coefficients were summed up and the percentage was taken to ascertain the correlation. It should be noted here that the correlation may be positive or negative but that should not matter. Hence all the coefficients are considered positive while summing.
- The percentage correlations obtained are the new Pollutant correlation (PC) weights. The vulnerability index is calculated by replacing the original weights with the PC weights.
- The model is validated using nitrate concentration and compared with other DRASTIC methodologies.
Table 1: Data Sources for the Models Parameter

<table>
<thead>
<tr>
<th>Data type</th>
<th>Sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>Meteorological Data</td>
<td>Kano state Water Board (KSWB. Technical Services Division)</td>
</tr>
<tr>
<td>Administrative Map of study area (Extent)</td>
<td>Global Administrative Areas GADM (2009)</td>
</tr>
<tr>
<td>Nitrate data (Sample wells)</td>
<td>Measured in the study area and location taken using geographical positioning system (GPS)</td>
</tr>
<tr>
<td>Topography</td>
<td>SRTM (30m) United States Geological Survey (USGS)</td>
</tr>
</tbody>
</table>

Fig. 1: Study area

Fig. 2: General methodology adopted in the study
RESULTS AND DISCUSSION

Nitrate data collected from 28 sample wells were interpolated using Kriging given by Equation (1).

\[ F(x,y) = \sum_{i=1}^{n} w_i f_i \tag{1} \]

where \( n \) is the number of scatter points in the set, \( f_i \) are the values of the scatter points and \( w_i \) are weights assigned to each scatter point.

The resulting nitrate concentration map is shown in Figure 3.

The rating of the DRASTIC parameter maps by [14] was maintained in this study (i.e. from 1-10). Hence when the rating were classified and interpolated were thus: For Depth to water (D), the whole area was found to be characterized by 1 (i.e. >30.48m) as shown in Figure 4A. Recharge (R) was found to inhabit the 1, 3, 6, 8 and 9 (Figure 4B). Aquifer media (A) rating was found to be 3, 4, 6 and 8 as shown in Figure 4C. The soil media (S) characterizing the study area were found to be 1, 2, 3, 5, 6, 8 and 9 as shown in Figure 4D. When the percentage slope of the study area was analyzed in defining the topography (T) the ranges obtained were 1, 3, 5, 9 and 10 as shown on Figure 4E. As for the impact of vadose zone (I) (Figure 4F) the material characterizing the area according to [14] were 1, 3, 4 and 6. The hydraulic conductivity (C) of the area when categorized was 1, 2, 4 and 6 as shown on Figure 4G.

Fig. 3: Nitrate concentration and map of study area

Pearson’s Product Movement Correlation (PPMC):

Scientists and engineers often collect data in order to define the nature of a relationship between two quantities in a view to studying the relationship between pair of experimental data. The experiment thus produces a collection of ordered pairs \((x_1, y_1), \ldots, (x_n, y_n)\) where \( n \) is the number of runs. Data that consist of ordered pairs are called bivariate data. When ordered pairs are plotted, they often tend to cluster around a straight line. The main question is commonly to determine how closely related the two quantities are to one other (Navidi 2010). The closeness of the association between the two variables is often defined by correlation coefficient. The correlation coefficient is usually represented by \( r \). To compute the correlation, the first step is to compute the deviations of the xs and ys, that is, \( x_i - \bar{x} \) and \( y_i - \bar{y} \) for each \( x \) and \( y \). The correlation coefficient is given by Equation (2).

\[ r = \frac{\sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2}} \tag{2} \]

The \( r \) is always between -1 and 1. Positive values of \( r \) implies that the least-squares line has a positive slope, meaning that greater values of one variable are associated with greater values of the other. The Negative values, however, indicate that the least-squares line has a negative slope, which means that greater values of one variable are associated with lesser values of the other. Values of the \( r \) close to 1 or to -1 indicate a strong linear relationship; values close to 0 indicate a weak linear relationship. \( r \) is equal to 1 (or to -1) only when the points in the scatterplot lie exactly on a straight line of positive (or negative) slope.

FRASTIC Index: The ‘FRASTIC Index’ was obtained by replacing the Depth to water (D) (Figure 4A) with the Fractured media (F) (Figure 4H) (the Fractured media was achieved by delineating the fractured zones within the study area) in the DRASTIC index evaluation method, while still maintaining the other 6 (R, A, S, T, I and C) parameter maps shown in Figure 4 (B, C, D, E, F and G) respectively. The choice of replacing D is based on the fact that the whole area is within the depth greater than 30.14m, hence characterized by common rating (i.e. 1). The vulnerability index was evaluated based on Equation (5) and the resulting FRASTIC index (FI) map is shown in Figure 7B. It is also worth noting that DRASTIC-PPMC and AHP approaches are not applicable to a parameter that indicates a common rating value throughout an entire area, because the correlation coefficient and vector of criteria computation for the two approaches are unattainable, respectively.
Fig. 4: Layer maps for interpolated parameters
Fig. 5: Created sampling points for data generation

Table 2: PPMC coefficients table for parameter vs. nitrate

<table>
<thead>
<tr>
<th>Parameters</th>
<th>Correlation coefficient (r)</th>
<th>Percentage Correlated weight (r) %</th>
<th>Normalized weight (Dw)</th>
</tr>
</thead>
<tbody>
<tr>
<td>F</td>
<td>0.006</td>
<td>1.562</td>
<td>0.359</td>
</tr>
<tr>
<td>R</td>
<td>0.090</td>
<td>22.607</td>
<td>5.200</td>
</tr>
<tr>
<td>A</td>
<td>0.049</td>
<td>12.426</td>
<td>2.858</td>
</tr>
<tr>
<td>S</td>
<td>0.095</td>
<td>23.874</td>
<td>5.491</td>
</tr>
<tr>
<td>T</td>
<td>0.058</td>
<td>14.651</td>
<td>3.370</td>
</tr>
<tr>
<td>I</td>
<td>0.066</td>
<td>16.576</td>
<td>3.813</td>
</tr>
<tr>
<td>C</td>
<td>0.033</td>
<td>8.334</td>
<td>1.917</td>
</tr>
</tbody>
</table>

\[ \sum_{i=1}^{n} r_i = 0.397 \quad \sum_{i=1}^{n} r_i = 100 \quad \sum_{i=1}^{n} Dw_i = 23 \]

**FRASTIC-PPMC:** Sampling points are created randomly over the entire study area. The points created were spread over the area so as to obtain a better representation of the area. Better result is expected if more points are created. In this study 202 points were generated as shown in Figure 5.

The points are then used to extract values from all the parameter maps shown in Figure 4 as well as that of nitrate (Figure 3). 'Extract values to point' spatial analyst tool in ArcGIS was used in the data extraction. Depth to water D was excluded because the whole area is characterized by a single digit (i.e. 1), hence, there may not be any correlation. The values extracted were then exported to Microsoft Excel. The \( r \) between the respective values for each parameter and nitrate was evaluated using Equation (4). The \( r \) values obtained for all the parameters are shown on Table 2. It is worth mentioning here that the negative sign for a negative correlation is ignored.

The \( r \) for all the parameters were summed up and the percentage of each is taken using Equation (3).

\[ r_w = \left( \frac{r}{\sum_{i=1}^{n} r_i} \right) \times 100 \]  

(3)
In order to maintain standard in vulnerability mapping and in a view to creating a favorable opportunity for indices comparison, the \( r \) values were normalized to obey with the generic DRASTIC (i.e. range from 23-230). The normalized DRASTIC weight is obtained based on Equation (4).

\[
D_W = \frac{r_w \times 23}{100} \quad (4)
\]

The new weights were then made to substitute the DRASTIC theoretical weight. The vulnerability index for the study area was evaluated based on Equation (5).

Equation (5) is the general equation for groundwater vulnerability assessment based on DRASTIC given by:

\[
DI = \sum_{j=1}^{7} (W_j \times R_j) \quad (5)
\]

where \( W \) is the weight, \( R \) is the rating while \( j \) represents the seven DRASTIC parameters.

The groundwater vulnerability map obtained when the PPMC weights were made to substitute the original DRASTIC (theoretical) weights (Table 3) is shown in Figure 6.

**SPSA and AHP Methodologies**

**Single Parameter Sensitivity Analysis (SPSA):** This technique is adopted in this study to estimate, singly, the influence of the DRASTIC parameter in each pixel or grid by relating its pre-assigned or ‘theoretical’ weight with the computed ‘real’ or ‘effective weight’. Equation (6) is used for the evaluation of effective weight of a parameter within a grid or pixel in a GIS based vulnerability assessment.

\[
W = (P \times r / V) \times 100 \quad (6)
\]

where \( W \) is the “effective weight” of an individual parameter, \( P \) and \( P_r \) are correspondingly the rating and the weight value of each parameter and \( V \) is the general vulnerability index.

The theoretical weights of the DRASTIC model (Table 3) were replaced with the effective weights obtained using the SPSA. The DRASTIC-SPSA and FRASTIC-SPSA maps obtained are shown on Figure 7C and 7D respectively.

**Analytic Hierarchy Process (AHP):** AHP generates weight for each parameter based on pairwise comparison of the DRASTIC parameters. The higher the obtained weight the more important the corresponding parameter. In AHP, scores are assigned to each option according to a pairwise comparison of the options based on that criterion. Higher score represents a better performance of the of the options vis-à-vis the criterion considered. The criteria weights and options scores are finally combined thereby ascertaining the global score (weighted sum of scores with respect to all criteria) for each option and resulting ranking. Four steps are engaged in implementing AHP (1) Vector of criteria weight computation (2) Matrix of option score computation (3) Options ranking (4) Checking for consistency.

\[
CI = \lambda - n/n - 1 \quad (7)
\]

The Consistency ratio (CR) is obtained using Equation (8).

\[
CR = CI/RI \quad (8)
\]
The general rule is that $RI \leq 0.1$ for a matrix to be consistent.

The DRASTIC–AHP and FRASTIC–AHP were obtained by replacing the original DRASTIC and FRASTIC weights with the new AHP weights (Table 3). The vulnerability indices were further re-evaluated using Equation (5). The resulting vulnerability index maps for DRASTIC–AHP and FRASTIC–AHP is shown in Figure 7E and 7F respectively.

The summary of the weights for different methods is shown on Table 3.

**Vulnerable Area Coverage by Model Type:** Figure 8 shows the vulnerable area coverage when the vulnerability index for each model is divided into 5 equal classes (i.e. very low, low, medium, high and very high). It has shown that the new developed F-PPMC model recorded highest area coverage for ‘very low’ (23.09%)
Fig. 8: Percentage vulnerable area coverage for each model

while D-AHP indicated the least area coverage of 6.17%. Generic DRASTIC recorded highest low class vulnerability area coverage of 45.86% as F-SPSA recorded the least (23.22%). As for medium vulnerability area coverage the maximum is seen on D-AHP (43.88%) and the minimum on F-AHP (20.2%).

For high vulnerability class area cover, F-AHP has the highest (22.59%) while the least record is that by DRASTIC (5.24%). The maximum area coverage according to very high vulnerability class is that attained in F-AHP model (13.77%) and the minimum is that of DRASTIC (2.52%).

When the average record for all the models was taken, it was found that very low is 16.04%, low is 30.98%, medium is 30.71%, high is 15.77% and very high is 6.51%.

Map Validation and Choice of Model Type: In validating the models, correlation between the nitrate and the models was established. This is achieved by comparing the vulnerability indices of each model and the nitrate contamination (calibrated see Figure 3) at a particular sample stations. 28 nitrate sampling point for all the models was utilized. Values observed at each sample station from individual vulnerability maps were analyzed using Pearson's correlation technique. The coefficients are shown on Table 4.

According to the validated maps, it has shown that the model that best suits the study area is the new developed DRASTIC-PPMC which correlated the best with the nitrate (i.e. 77%). Nitrate indicates the actual groundwater contamination trend of the study area.

The percentage total area coverage for each class based has shown that very low covers only 10%, low 26%, medium is 38% and high 19.9% and very high 7%, as depicted in Figure 10. Generally, the study area can be said to be characterized very low to medium vulnerability since the very high is negligible.

<table>
<thead>
<tr>
<th>Pearson's correlation coefficient</th>
<th>Sample stations</th>
<th>Factor</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.73</td>
<td>28</td>
<td>Nitrate</td>
</tr>
<tr>
<td>0.65</td>
<td></td>
<td>DRASTIC</td>
</tr>
<tr>
<td>0.75</td>
<td></td>
<td>FRASTIC</td>
</tr>
<tr>
<td>0.75</td>
<td></td>
<td>DRASTIC-SPSA</td>
</tr>
<tr>
<td>0.75</td>
<td></td>
<td>DRASTIC-SPSA</td>
</tr>
<tr>
<td>0.74</td>
<td></td>
<td>DRASTIC-AHP</td>
</tr>
<tr>
<td>0.63</td>
<td></td>
<td>FRASTIC-AHP</td>
</tr>
<tr>
<td>0.77</td>
<td></td>
<td>FRASTIC-PPMC</td>
</tr>
</tbody>
</table>

In a study by [56] the DRASTIC model was validated with nitrate concentration and a coefficient of 0.42 was realized. In that of [55] the nitrate correlation (Pearson’s) level with the modified DRASTIC-SPSA a coefficient of 0.58 was achieved. [47] utilizes similar technique using 27 locations and obtained a relationship of 0.44 and 0.82 for the generic DRASTIC and modified DRASTIC-SPSA models respectively. [54] also had the opinion that the parameter ratings were modified based on the relationship between the parameters and nitrate concentration could be used as a modifying parameter with considerable
improvement in the resulting index whereby the rating were modified and significant results were achieved as when validated the initial correlation was 0.23 and after modification, it improved to 0.68. Similarly, in a study by [46], the DRASTIC ratings were modified using frequency ratio method (FR) where in the original DRASTIC vulnerability correlation was 0.37. After optimization using FR-DRASTIC, the correlation improved to 0.75. Spearman rank correlation was (using 24 points) was adopted to modify by revising the rating scale of each index in a study by [3], the method improved the models validity from 0.4098 to 0.6698. Although there are several methodologies employed using correlation between nitrate and DRASTIC parameters to develop weights or ratings, this method has proven to be a direct and more consistent method.

CONCLUSION

Protection of groundwater from pollution is generally a global political decision. These decisions are essentially supported by DRASTIC analysis-being the most popular method of evaluating the intrinsic pollution vulnerability of groundwater. However, because of the spatial variability around the world, the model has been exposed to various transformations. Especially the weights assigned to the parameters constituting the model, for their enormous effect towards achieving final vulnerability result. Popular among the DRASTIC weights transformation technique are the single parameter sensitivity analysis (SPSA) and analytic hierarchy process (AHP). While these methods are proven effective in some studies, they are contrarily ineffective in others. However, the new approach developed using Pearson’s product moving correlation (PPMC) in this study (i.e. DRASTIC-PPMC) had proven to be the most effective groundwater pollution vulnerability assessment approach for Kano Nigeria based on the seven (7) models employed. Although there are several methodologies employed using correlation between nitrate and DRASTIC parameters to develop weights or ratings, this method has proven to be a direct and more consistent method. Furthermore, since weight assigned to a parameter is directly proportional its level of input on attaining nitrate contamination, it is likely to be most suitable for virtually all areas as the groundwater vulnerability map is generated based on the pollution pattern of any area under study.

When the area characterized based on quantile classification it was established that the study area has a very low vulnerable area cover of 10% predominantly in the eastern parts, 26% low within the east and western parts, 38% medium, mainly in the north and eastern parts. 19% high and only 7% very high vulnerability which is principally in the extreme north.

ACKNOWLEDGEMENTS

The support provided by the management of Kano Agricultural Rural Development Authority (KNARDA), Kano State Water Board (KSWB) Technical Services Division, Challawa Workshop for the data are at this moment acknowledged. Moreover, University Putra Malaysia under the Research University Grant Scheme (RUGS), 05-02-12-1874RU, 9344400 is acknowledged.

REFERENCES


