Geostatistics approach with indicator kriging for assessing groundwater vulnerability to Nitrate contamination

Rohazaini Muhammad Jamil[1], Mohamad Nor Said[2] and Mohd Nadzri Md Reba[3]

Department of Geoinformatics, Faculty of Geoinformation and Real Estate,
Universiti Teknologi Malaysia, 81310 UTM Skudai, Johor, Malaysia.
[1]rohazaini@umk.edu.my, [2]m.nor@utm.my, [3]nadzri@utm.my

Abstract

In Malaysia, the demand for water has increased tremendously and groundwater has been identified as one of the alternative to new water sources. In north eastern state of Kelantan, Malaysia, almost 70 percents of people consumes groundwater in their daily lives. However, due to uncontrolled development and human activities, groundwater is subjected to pollution and geochemical contamination. This paper is intended to assess groundwater vulnerability to nitrate (NO_3) contamination using geostatistical approach with indicator kriging instead of using the typical DRASTIC (Depth, Recharge, Aquifer, Soil, Topography, Influence of the vadose zone and Conductivity) model. Geostatistical approach has the capability to predict and understand the geochemical concentration pattern as an alternative to the DRASTIC model which is unable to provide the measurable information of geochemical characteristics. The results of the modelling are produced and presented in the form of probability map using indicator kriging interpolation. The finding of the study can be used to identify the probability of geochemical concentration exceeding the specified threshold.

Keywords: Groundwater, geostatistics, indicator kriging, DRASTIC model

1. Introduction

The shallow aquifer system in the Kelantan river delta constitutes an important source of water not only for public water supply, but also for domestic and agricultural purposes. Being shallow and relatively unprotected, the aquifer is generally exposed to higher risk of contamination due to anthropogenic activities at the surface. In Kelantan, the climate changes are obvious and produce significant impacts on groundwater levels which also contribute towards the contamination of groundwater in the shallow aquifers. Nitrate, being one of the major contaminants in groundwater may be hazardous when it is consumed excessively. Studies have shown that high nitrate levels are associated with diseases such as methemoglobinemia, gastric cancer, non-Hodgkin's lymphoma and diarrhea (Mohamed Azwan et al., 2010). The typical standard value for nitrate concentration in groundwater based on Department Of Environment (DOE) standard in Malaysia is 10mg/L.

Groundwater quality monitoring is an essential step in understanding groundwater phenomenon. This step involves collecting samples and carrying out analysis in the lab, which makes it costly. Increasing the number of sampling points will probably increase the amount of information obtained from the network. There are two approaches in groundwater monitoring: the hydrogeological approach and statistical approach using kriging interpolation. The widely used hydrogeological approach is groundwater vulnerability assessment as the term "vulnerability" was defined as the degree of protection that the natural environment provides against the spread of pollution in groundwater.

The concept of vulnerability of groundwater to contamination was introduced by Margat (1968). But, this definition has gradually changed, e.g. Hrkal (2001) defines groundwater vulnerability as the tendency and likelihood for general contaminants to reach the water table after introduction at the ground surface. "Vulnerability" is the degree of which human or environmental systems are likely to experience harm due to perturbation or stress, and can be identified for a specific system, hazard, or group of hazards (cited by Liggett and Talwar, 2009 in Popescu et al. 2008). Groundwater is vulnerable to contamination by anthropological activities and it is very difficult to remediate once contaminated. To properly manage, protect and monitor the resource, it is therefore important to determine areas with more aspects of vulnerable to contamination. Assessments of the vulnerability often results in a map of areas where the resource is vulnerable to contamination due to the surface activities and finally optimize the groundwater quality monitoring.

Groundwater vulnerability assessments are meant to synthesizing complex hydrogeologic information into a feasible map to be used by planners, decision and policy makers, geoscientists, and the public. The vulnerability maps is useful for many aspects of water management, including hot spot areas (likely to be contaminated) for monitoring, protection, and further investigation. Generally, vulnerability assessments are categorized as (i) index (and overlay) methods; (ii) statistical methods; or (iii) process methods (Focazio et al. 2002). The examples of models for each of these methods are shown in Table 1.

Table 1: Examples of each of the three methods of vulnerability assessment (Liggett and Talwar, 2009)

Methods	Models (examples)
Index and Overlay methods	DRASTIC
Statistical Methods	Logistic Regression
Process Methods	Numerical Models

Index and overlay method is widely used by the hydrogeologists and engineers for assessing groundwater vulnerability. This popular model which is called DRASTIC model (Liggett and Talwar, 2009) has been chosen to be applied in this study. Another practice of assessing vulnerability is by using statistical method. This method is based on observations (i.e. not on expert opinion) and anthropogenic factors that are easy to update as new available information. But, this method is less popular because it needs experts and difficult to be interpreted by the non-mathematical person. Meanwhile, the third model approach is process based method but this method requires additional interpretation by the water resource decision maker and scientists' consultation in order to meet certain policy and management objectives. Basically the numerical model is applied in this process based method.

As compared to these three methods, the DRASTIC model (by means of category of overlay and index method) is widely used in identifying areas at which groundwater supplies or well are most susceptible to contamination (Aller et al., 1987 and Sener et al., 2009). This method is the most popular groundwater vulnerability assessment as it is inexpensive, straightforward, and relying on data that are commonly available or estimated, and producing end-products that are easy to be interpreted and incorporate into the decision-making process (Focazio et al., 2002). Despite the fact that DRASTIC is popular, this model has yet to be scrutinized in order to find the potential improvement as its limitation has been revealed by the past studies. Evans and Maidment (1995) claimed that the DRASTIC model is lack of methodological foundation as the introduction of contaminant loading in the DRASTIC model unfortunately contributes very little to the correlation of the actual contamination occurrence with the derived scores.

Instead of these limitations, a study from Centre of Research in Water Resources, Texas (Evans and Maidment, 1995) claimed that a statistical approach by adding GIS approach are better to be applied compared to DRASTIC model alone. It is also proven can be applied straightforward as an index or overlay method, and gives more defensible foundation. However, this method is the least popular vulnerability assessment in the literature and it is normally used as a test for other methods. Geostatistics is a branch of science that applies statistical methods to spatial interpolation in GIS environment. Although geostatistics was developed independently of GIS, it has become an integral part of GIS. Without a computer and GIS mapping ability, it wouldn't be known outside a small group of geostatistical experts. Geostatistics are considered to be one of the most sophisticated spatial interpolation methods. The method is commonly used in many disciplines such as geology, engineering, hydrology, geography, ecology, urban studies, and medical geography. Geostatistics are closely related to statistics and GIS.

Geostatistical method such as kriging interpolation is a ubiquitous approach that quantitatively estimates the distribution of water quality parameters. However there is no work related to the application of geostatistical method in the groundwater vulnerability assessment has neither been published nor verified (GAO, 1992). Therefore, the main objective of this study is to assess groundwater to contamination by using the geostatistical method to model the continuous surface of geochemical concentration and finally to determine which spot (i.e. well) has exceeded the threshold value of geochemical concentration. This paper discusses the typical issues in DRASTIC model and introduces geostatistics as an alternative approach.

2. Groundwater vulnerability assessment based on DRASTIC model

DRASTIC has been used for mapping aquifer vulnerability in porous aquifers (Aller et al. 1987). The DRASTIC method is derived from ratings and weights associated with the seven parameters namely the depth to groundwater (D), the net recharge (R), the aquifer media (A), the soil media (S), topography (T), influence of the vadose zone (I) and the hydraulic conductivity (C). Each parameter is subdivided into ranges with different ratings assigned in a scale of 1 to 10 (least to highest contamination potential). A higher DRASTIC index shows greater groundwater pollution vulnerability (Sener et al., 2009).

Yet, due to its limitation, some studies have applied some modifications and improvements to the DRASTIC procedures (Thirumalaivasan et al., 2003; Dixon, 2005; Panagopoulus et al., 2006). According to Rosen (1994), some important scientific parameters such as sorption capacity, travel time and dillution are not taken into account directly. Besides, it is very difficult to test the accuracy of the model because it requires properties of pollutant to be assumed as a model and deposited for all over the test area at a uniform concentration and for a considerable period of years in order to allow the hydrogeological setting to respond.

Due to the fact that DRASTIC approach is unable to provide a measurable information of chemical characteristics (which is very important to the environmentalists, scientists or hydrogeologists), thus there is a need to introduce an alternative method to cope with these drawbacks.

2.1 Geostatistics and indicator kriging

The information of chemical characteristics is very important to the environmentalists, scientists or hydrogeologists to quantitatively define the characteristics of surface and volumes of regionalized variables which have been sampled only at discrete points. Geostatistical approach has been identified to be an alternative way in assessing groundwater vulnerability to contamination. The geostatistics features a statistical model to appropriately estimate the continuous area (i.e. groundwater contamination phenomenon) in a discrete form.

The principles of geostatistics have been applied to a variety of areas in geology and other scientific disciplines (Nelson et al., 1999, Hamad, 2009). The major concern is that of defining quantitatively the characteristics of surface and volumes of regionalised variables which have been sampled only at discrete points. This has raised a question on how possible one to estimate the value of a variable in a site if the scientists and engineers are provided with a set of sampled data collected at specific location within an area under investigation.

Therefore, Geographic Information Systems (GIS) coupled with geostatistics application have the capability to help in virtually display and present the assessment results. The geostatistics is a handy tool that works with continuous data such as rainfall, temperature, geology, or soils to create a surface (probability and prediction surfaces). Kriging is a powerful geostatistical technique that applies interpolation and prediction at the unsampled location and utilises spatial relationships to give a better and accurate predictions. It was named after D.G Krige from South Africa (Sahoo, 2003). The kriging interpolator is considered the most sophisticated and accurate way to determine the intensity of a phenomenon at unmeasured locations and the weights surrounding measured values are based not only on the distance between measured points and the prediction location but also on the overall spatial arrangement of the measured points. This techniques provides a linear estimation as a function of the variable values at nearby locations. Kriging has two main advantages with respect to other linear estimators: a) the weights used in the estimation are determined as a function of both the structural distance (variogram) between the data and the location to be estimated, and the structural distance between any pair of two data, and b) the estimation is accompanied by an uncertainty quantification, the kriging error variance.

Meanwhile, indicator kriging uses threshold to create binary data (0 or 1 value), and then uses ordinary kriging to make spatial predictions based on the indicator data. The data points need to be sampled from a phenomenon that is continuous in space. Predictions using indicator kriging are interpreted as the probability of exceeding (or being below) the specified threshold. The validity of indicator kriging relies heavily on the assumption of stationarity and it should not be used with data having a trend. Since the indicator variables are 0 or 1, the interpolations will be between 0 and 1, and predictions from indicator kriging can be interpreted as probabilities of the variable being 1 or being in the class that is indicated by 1. If a threshold was used to create the indicator variable, then the resulting interpolation map would show the probabilities of exceeding (or being below) the threshold. The resultant probability map will indicate where the interpolated values exceed a specified threshold.

3. Methodology

There are two major phases of methodology in this study, (1) data collection and (2) modelling using geostatistical method. Data collection phase is illustrated in Figure 1.

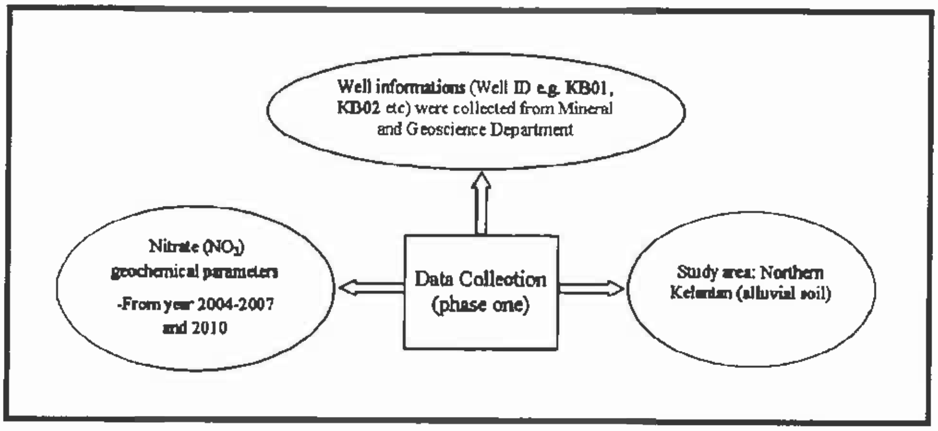

Figure 1: Data collection phase.

This study was conducted at the north eastern state of Kelantan, which is the most populated area with groundwater wells mostly found. It is a potential spot in which most of the groundwater sampling and monitoring activities are systematically carried out by the Department of Mineral and Geoscience (JMG). Nitrate geochemical concentration parameter has been selected to be modelled as it shows a significant fluctuation of concentration level from year 2004 till year 2007 (Rohazaini, 2008). The latest data in year 2010 has also been collected from JMG to study the probability of nitrate geochemical concentration exceeding the threshold (DOE standard).

The second phase is data modelling using geostatistical approach as shown in Figure 2. In this study, Geostatistical Analyst of ArcGIS 9.2 is used to model the groundwater contamination maps. Instead of providing various interpolation techniques, the geostatistical feature provides (prior to mapping) Exploratory Spatial Data Analysis (ESDA) tools that help to assess the statistical properties of the data. In the exploratory spatial data analysis, variety of output map types (prediction, error of prediction, probability and quantile) are created using kriging and co-kriging variants (ordinary, simple, universal, indicator, probability and disjunctive) and auxillary tools (data transformation, declustering, and detrending). Indicator kriging interpolation has been chosen to be applied as this method is more suitable to produce the probability map. The purpose of creating the probability map is used to locate and identify the risky area and well which exceeds the critical threshold or standards of the DOE.

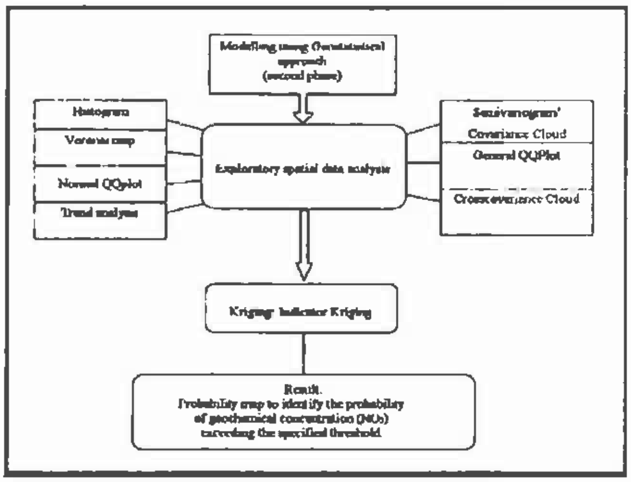

Figure 2: Modelling phase

4. Results and discussion

In the decision-making process, care must be taken in using a map of predicted contaminated wells for identifying unsafe areas because it is necessary to understand the uncertainty of the predictions. For example, suppose the critical threshold of DOE standard value is 10 mg/litre for nitrate and any locations exceed this value will be determined. As Geostatistical Analyst provides a number of methods that can perform this task, indicator kriging have been chosen as it is the simplest available model to be applied. This technique does not require the dataset to conform to a particular distribution. The data values are transformed to a series of 0s and 1s according to whether the values of the data are below or above a threshold. If a threshold of 10 mg/litre is used, any value below this threshold will be assigned a value of 0, whereas the values above the threshold will be assigned a value of 1. Indicator kriging then uses a semivariogram model that is calculated from the 0–1 dataset.

The resultant of nitrate geochemical concentration probability maps show that the geostatistics has a great potential in determining the percentage of nitrate concentration exceed the DOE standard within the study area. The maps have been modelled started from year 2004 till 2007 and the latest one is from year 2010. As referred to Figure 3, it shows the probability map of nitrate exceeding the DOE standard in year 2010. The excessive level of contamination is presented in darkest purple contour and being determined as risky area that is vulnerable to contamination. Meanwhile, Figure 4 and 5 show in year 2010, five monitoring well locations have been affected by nitrate contamination and having a 42.4% possibility of exceeding the DOE standards as the calculation value for the darkest purple contour is 0.424474. The sequence or pattern of the resultant maps from year 2004 till 2007 can also be spatially viewed as shown in Figure 7. The exceedence of specified threshold (DOE standards) can be considered as that particular area is risky and vulnerable to contamination. This analysis is very important to the regulatory bodies or Department of Environment to make decisions in monitoring water quality in the study area.

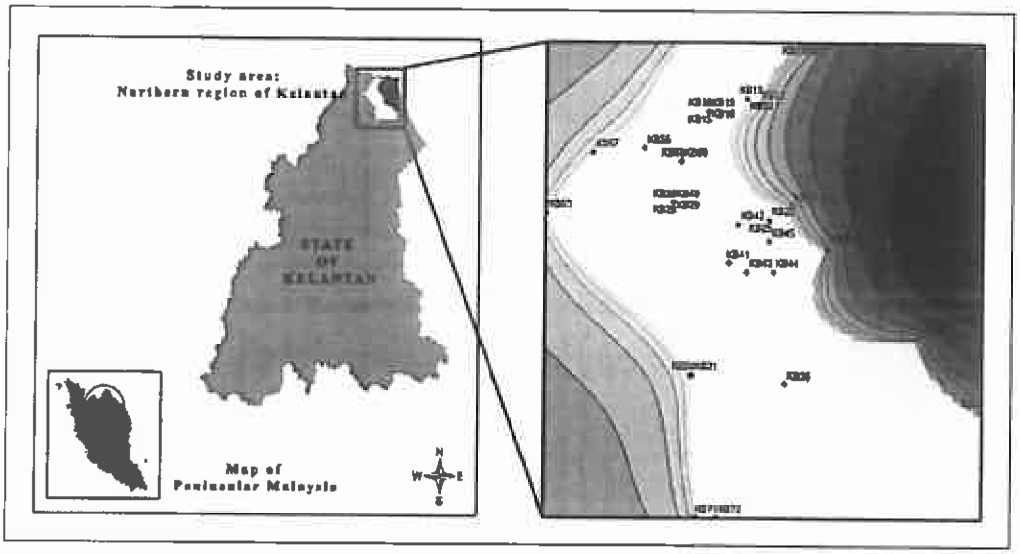

Figure 3: The excessive level of nitrate contamination in year 2010 is presented in the darkest purple contour.

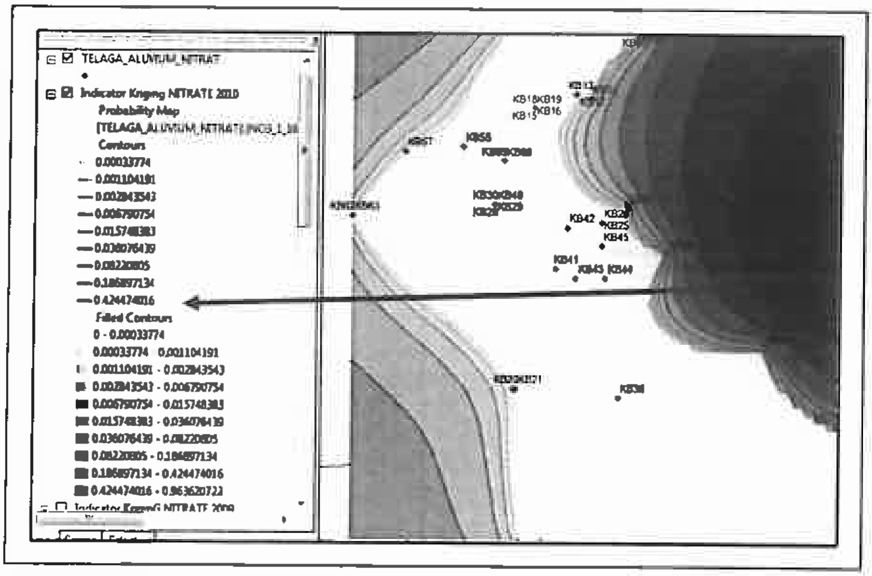

Figure 4: The darkest purple contour shows the probability of nitrate concentration exceeds the DOE standards in 2010

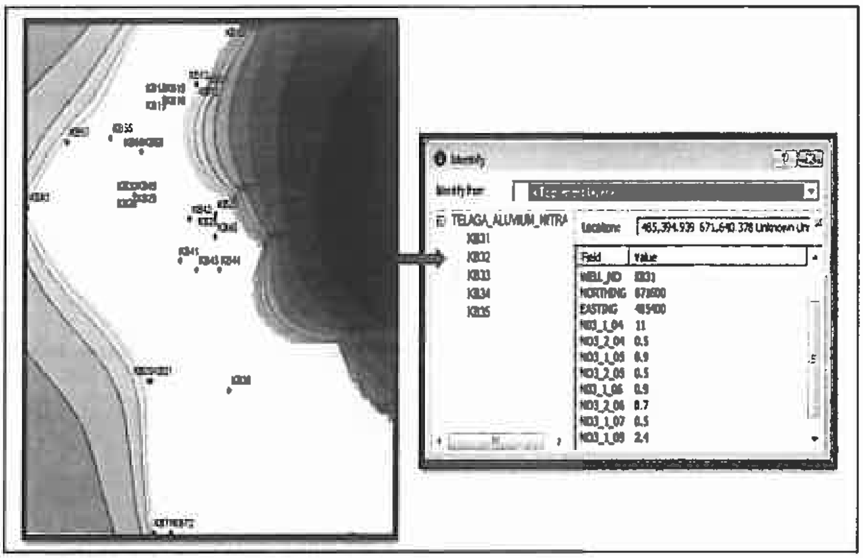

Figure 5: The resultant map shows the information of which location/well that is having a 42.4% possibility of exceeding nitrate DOE standards

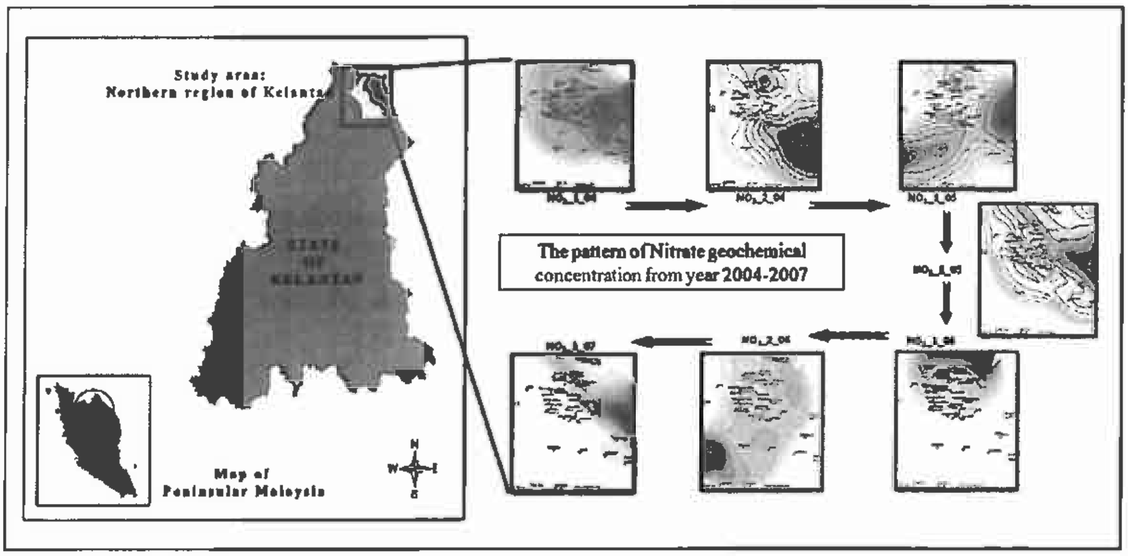

Figure 6: The pattern of nitrate geochemical concentration from year 2004 till 2007

5. Conclusion

Spatially related statistics or geostatistics can be applied in hydrogeology field especially in groundwater vulnerability to contamination issues. This study is useful in developing groundwater protection strategies and for optimal monitoring, water management decision makers, as well as regulatory bodies' attempts for the regions in which the prevention of groundwater contamination seems most needed. For example, it will help planners and decision makers on proper land use and water resource management such as to avoid high risk areas when locating a site of pollution potential or selecting the areas for waste disposal and industrial sites. In future, geostatistics is expected be applied spatially in natural resources evaluation for practitioners in such diverse fields as soil science, mining, petroleum, remote sensing, hydrogeology, and the environmental sciences.

Acknowledgements

The authors are thankful to the Mineral and Geoscience Department of Kota Bharu Kelantan for their help and priceless support in ensuring the success of this study.

References

Aller, L., Bennett., Lehr, J.H., Petty, R.J., Hackett, G., (1987), DRASTIC: A Standardised System For Evaluating Ground Water Pollution Potential Using Hydrogeologic Settings, *US Environmental Protection Agency Report (EPA/600/2-87/035)*, Robert S. Kerr Environmental Research Laboratory, 455 pp.

Dixon, B., (2005), Groundwater Vulnerability Mapping: A GIS And Fuzzy Rule Based Integrated Tool, *Journal Of Applied Geography* 25(2005) 237-347.

Evans, T.A, and Maidment, D.R., (1995), A Spatial And Statistical Assessment Of The Vulnerability Of Texas Groundwater To Nitrate Contamination, *Centre For Research In Water Resources*, The University of Texas, Austin. Retrieved from http://civil.ce.utexas.edu/centers/crwr/reports/online.html (Last accessed 27 January 2011).

Focazio, M.J., T.E. Reilly, M.G. Rupert, and D.R. Helsel., (2002), Assessing Groundwater Vulnerability To Contamination: Providing Scientifically Defensible Information For Decision Makers, US Department of Interior and US Geological Survey, Reston, *VA. US Geological Survey Circular No. 1224.*

GAO (General Accounting Office) (1992). Goundwater Protection: Validity and Feasibility of EPA's Differential Protection Strategy. *(GAO/PEMD-93-6) Washington, D.C.: General Accounting Office.* Retrieved from http://www.legistorm.com/showFile/L2xzX3Njb3JlL 2dhby9wZGYvMTk5Mi8xMg==/ful22507.pdf (Last accessed 26 January 2011).

Hamad, S., (2009), Geostatistical Analysis Of Groundwater Levels In The South Al Jabar Al Akhdar Area using GIS, Proceedings *Of GIS Ostrava 2009, January, 25-28.* Retrieved from http:www.gis.usb.cz/GIS_Ostrava/GIS_Ova_2009/Sbornik/Lists/Papers/005.pdf (Last accessed 5 February 2011).

Hrkal, Z., (2001), Vulnerability Of Groundwater To Acid Deposition, Jiserke Mountains, Northern Czech Republic: Construction And Reliability Of A GIS-Based Vulnerability Map, *Journal Of Hydrogeology*, 9(2001), 348-357.

Liggett J.E., and Talwar, S., (2009), Groundwater Vulnerability Assessments And Integrated Water Resource Management, Streamline Watershed Management, *Vol. 13/no.1 Bulletin Fall 2009*, article 4, 18-29 .

Margat, J., (1968), Groundwater Vulnerability Maps, *Conception Estimation Mapping (in French)*, EEC Institute European de l'Eau, Paris, France, p11.

Mohamed Azwan, M.Z., Mohd Kamil, Y., Hazilia, H., and Suhani, N., (2010), Nitrate-Nitrogen Concentration Variationin Groundwater Flow In A Paddy Field, *The Institution of Engineers Malaysia Journal* (Vol. 71, No.4, December 2010), 1-9.

Nelson, M.R., Orum, T.V., Nadeem, A., and Garcia, R.J., (1999), Applications Of Geographic Information Systems And Geostatistics In Plant Disease Epidemiology And Management, *The American Phytopathological Society, Plant Disease*, 4(83), 308-319.

Panagopoulus, G.P., Antonakos, A.K., and Lambrakis N.J., (2006), Optimizing Of The DRASTIC Method For Groundwater Vulnerability Assessment Via The Use Of Simple Statistical Methods And GIS, *Hydrogeology Journal* 14(2006), 894-911.

Popescu, I.C., N. Gardin, S. Brouyére, and A. Dassargues., (2008), Groundwater Vulnerability Assessment Using Physically-Based Modelling: From Challenges To Pragmatic Solutions, in *ModelCARE 2007 Proceedings*, Calibration and Reliability in Groundwater Modelling, J.C. Refsgaard, K. Kovar, E. Haarder, and E. Nygaard (editors), Denmark, IAHS Publication No. 320.

Rohazaini, M.J., (2008), Permodelan Pencemaran Geokimia Air Tanah Dalam Lapisan Aluvium Di Negeri Kelantan Menggunakan Geostatistical Analysts (Arcgis 9.2), *Bachelor of Science Geoinformatics Thesis*, Faculty of Engineering and Science Geoinformation, Universiti Teknologi Malaysia.

Rosen, L., (1994), A Study Of The DRASTIC Methodology With Emphasis On Swedish Conditions, *Ground Water* 32(1994), 278–285.

Sahoo, R.N., (2003), Geostatistics In Geoinformatics For Managing Spatial Variability, *Indian Agricultural Research Institute*, New Delhi. Retrieved from http:www.iasri.res.in/ebook/EB_SMAR/e-book_pdf%20files/manual%20IV/6-Geostatistics.pdf (Last accessed 5 February 2011).

Sener, E., Sener, S., and Davraz, A., (2009), Assessment Of Aquifer Vulnerability Based On GIS And DRASTIC Methods: A Case Study Of The Senirkent-Uluborlu Basin (Isparta, Turkey), *Journal of Hydrogeology* 17(2009), 2023-2035.

Thirumalaivasan, D., Karmegam, M., and Venugopal, K., (2003), AHP-DRASTIC: Software For Spesific Aquifer Vulnerability Assessment Using DRASTIC Model And GIS, *Journal of Modelling and Software* 18(2003), 645-656.