

Assessing groundwater quality conditions in the upper Litani River Basin utilizing geostatistical analysis tools

H. Assaf¹ & M. Saadeh²

¹*Department of Civil and Environmental Engineering, American University of Beirut, Lebanon*

²*Environmental Unit, Litani River Authority, Lebanon*

Abstract

Located in a region known for its scarce water resources, The Litani River Basin, totally contained within Lebanon's borders, is endowed with significant water resources that have been at the centre of Lebanon's water supply, agriculture and hydroelectric development plans. However, many years of civil strife and foreign occupation compounded with proliferation of uncontrolled urban, agricultural and industrial growth have taken their toll on the quality of the basin's surface water and potentially its groundwater quality conditions.

This paper reports on findings from a study with the objective of assessing groundwater quality conditions in the Upper Litani Basin (ULB) based on data collected through an extensive water quality survey funded by the USAID as a component of its initiative to support local government efforts to manage water quality conditions in the country. The data is composed of two sampling sets representing summer and winter conditions and include several key water quality indicators. The paper focuses on the methodology and results of a geostatistical analysis of the sampled groundwater nitrate concentrations. The analysis produced two sets of georeferenced nitrate concentration and probability of exceedance maps representing winter and summer conditions. The results indicate a significant, widespread and persistent nitrates contamination of groundwater in the ULB. Nitrate levels in groundwater exceed standard limits for drinking water in many parts of the basin.

Keywords: water quality, water, geostatistics, Kriging, water management, groundwater, Lebanon, nitrates, Litani, contamination.



1 Introduction

Situated in the world's most water-deprived region, Lebanon stands out as a country with relatively rich water resources, mainly originate from heavy winter orthographic precipitation intercepted by sharply rising mountain ranges that parallel the Mediterranean coast. Rising near the ancient city of Baalbek, 85 km east of the capital Beirut, the Litani River flows 140 kms in south-westerly, southerly and westerly directions, to meet the Mediterranean 70 kms south of Beirut. The river drains over fifth of Lebanon's total area of 10,500 km² and is totally contained within its border making it the country's most important water resource. In the late 1950s, a major hydroelectric system was constructed to tap the 800 m head between the Mediterranean and the river site near the town of Qaraoun, 70 kms downstream from Baalbek. The development involved constructing the Qaraoun dam and diverting the river via a series of ponds and tunnels and through 3 hydroelectric plants to empty into the Mediterranean 30 kms north of the river's original outlet. The development has resulted in the hydrological separation between the subbasin above Qaraoun lake known as the Upper Litani Basin (ULB) and the Litani's lower reaches.

The ULB is home to over 500,000 inhabitants mostly engaged in agricultural activities, and food processing and tourism industries. Producing the bulk of Lebanon's food output, agriculture in the ULB relies on pumped water from surface water and groundwater resources during the rainless summer season. However, relentless releases of untreated domestic, industrial and agricultural wastes into the river have seriously degraded surface water quality conditions to hazardous levels. Water quality is expected to deteriorate even more rapidly if significant control measures are not implemented immediately (Assaf and Saadeh [1]).

Intensive application of fertilizers, releases of untreated wastewater to open areas, ditches and septic tanks and the common practice of dumping solid waste, including animal carcasses and industrial waste, in many parts of the basin are suspected to have undermined groundwater quality conditions in the ULB. A comprehensive study is undertaken to assess the level and extent of suspected nutrient and bacterial contamination of groundwater in the ULB to support development of policy options to manage the rising risk of contamination to human health and environment. The current paper focuses on the results from the assessment study with respect to groundwater nitrate contamination. The analysis is based on data collected through an extensive water quality survey funded by the USAID as a component of its initiative to support local government efforts to manage water quality conditions in the country (BAMAS [2]). The data includes two sets of samples representing winter (February/March) and summer (June) conditions collected in 2005 from 60 wells across the basin (fig. 1). The data shows significantly elevated nitrates (NO₃) concentrations (mg/L), especially in the winter.

At higher levels, nitrates constitutes a serious health risk as it disintegrates in the body into nitrite, which hampers oxygen transfer by binding with haemoglobin and leading to Methemoglobinemia, which is particularly life



threatening to infants (Ward *et al* [9]). Also nitrate is a precursor to the development of the genotoxic N-nitroso compounds (NOC), which are known animal carcinogens (Ward *et al* [9]).

The study utilizes geostatistical analysis tools to produce maps showing spatial distribution and probability of exceedance of groundwater nitrate concentrations.

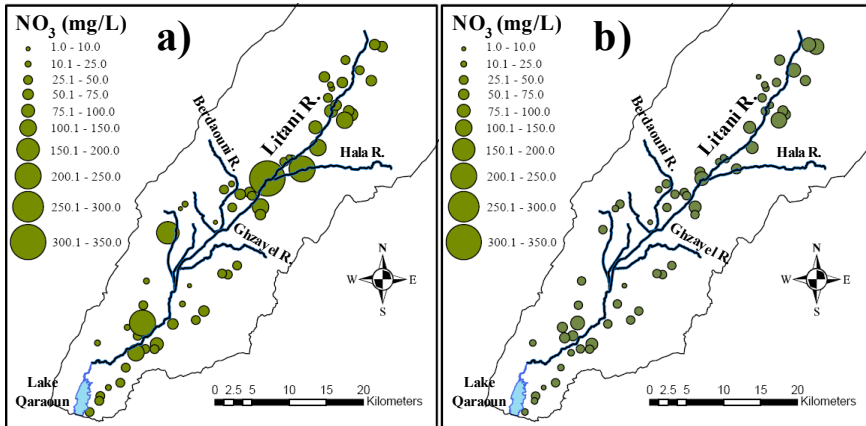


Figure 1: Measured groundwater nitrate concentrations in the ULB: a) winter, b) summer.

2 Methodology

Critical to the analysis of spatial information as the one addressed in this paper, is the ability to reliably estimate and present its spatial distribution from point data. A powerful approach to achieve this objective is the one advanced by geostatistics, where a continuous surface representing a given variable is calculated from point data based on the potential presence of correlation among data points as a function of the modulus and direction of vector separating them. Known as spatial continuity, this relationship is an important characteristic of spatial data that can provide insight into the physical nature of the phenomena under study.

Formally, geostatistics is a branch of applied statistics that deals with detection, modelling and estimation of spatial patterns (Rossi *et al* [8]). It was originally adopted by the mining industry to estimate mineral reserves based on ore samples. It was subsequently applied in a diverse range of problems including ecological studies, air pollution, groundwater and soil contamination. The trend accelerated with the advent of increasingly powerful computer systems and sophisticated geostatistical analysis tools in particular those integrated within Geographic Information System (GIS) applications.

2.1 Theoretical background

Geostatistics is based on random theory principles. Consider a given spatial medium, e.g. a groundwater contaminant plume, where n measurements, $z(x_1), \dots, z(x_n)$, of the variable under consideration are collected at points x_1, \dots, x_n . Geostatistics is based on the notion that the set of measurements represents a single realization of the random function $Z(x)$ for all possible values in the medium. Values at unmeasured locations can be best estimated based on the conditional expectation (Cooper and Istok [5]):

$$E \left[Z(x_0) \mid Z(x_1), \dots, Z(x_n) \right] \quad (1)$$

where x_0 is a point where no measurement is available.

Geostatistics, or more specifically linear geostatistics, relies on two assumptions to solve eqn. (1) numerically. $Z(x)$ is assumed to be normally distributed, thus only two statistical moments (the mean and the variance) need to be specified. The probability distribution is also assumed to be stationary over the medium. Each measurement can be then considered an individual realization of $Z(x)$. Consequently the set of measurements can be used to calculate the mean and variance of $Z(x)$.

Kriging refers to the process of estimating variable values at unmeasured locations based on eqn. (1). Given the assumption of normality, the value at an unmeasured location, $Z(x_0)$, is estimated as a weighted average of measured values as follows:

$$Z^*(x_0) = \sum_{i=1}^n \lambda_i Z(x_i) \quad (2)$$

where $\lambda_1, \dots, \lambda_n$ are a set of weights calculated based on solving the following Kriging system designed to produce the best linear unbiased estimate:

$$\sum_{j=1}^n \lambda_j \gamma(x_i, x_j) + \mu = \gamma(x_0, x_i); \quad i = 1, \dots, n \quad (3)$$

$$\sum_{j=1}^n \lambda_j = 1 \quad (4)$$

where $\gamma(x_i, x_j)$ is the semivariogram defined as:

$$\gamma(\mathbf{h}) = 1/2 E \left\{ \left[Z(x + \mathbf{h}) - Z(x) \right]^2 \right\} \quad (5)$$

The semivariogram is a measure of spatial continuity and depends only on the vector \mathbf{h} separating a given pair of measurements, and not the positions of these measurement. A smaller $\gamma(\mathbf{h})$ indicates a higher correlation and vice versa. Assuming isotropic conditions, i.e. ignoring \mathbf{h} direction, the semivariogram is estimated by the experimental semivariogram $\gamma^*(h)$ calculated as follows:



$$\gamma^*(h) = \left[\frac{1}{2N(h)} \right] \sum_{i=1}^{N(h)} [z(x_{i+h}) - z(x_i)]^2 \quad (6)$$

where $N(h)$ represents the number of pairs of measurement points separated by a distance h . $\gamma^*(h)$ is not used directly in the Kriging system (eqns. (3) and (4)) since it is not available for all possible h values. Instead a mathematical model is fitted to the experimental semivariogram in a process known as structural analysis. Any mathematical model can be used given that it is positive definite, a property that is imposed to guarantee a unique solution of the Kriging system (Journel and Huijbregts [7]). Fig. 2 shows the main components of a semivariogram model. $\gamma(0)$, which should theoretically be equal to zero, is known as the nugget. A nonzero nugget reflects sampling and analysis errors. The range is the distance beyond which $\gamma(h)$ does not change significantly. A longer range indicates a stronger spatial continuity. The sill is $\gamma(h = \text{range})$. The partial sill is the difference between the sill and the nugget.

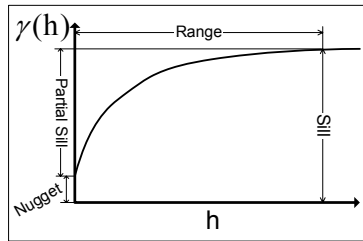


Figure 2: A generic semivariogram model.

2.2 Geostatistical analysis approach

The application of geostatistics involves three main steps: 1) exploration of the data to characterize its spatial continuity and assess its suitability for geostatistical analysis 2) structural analysis to develop a semivariogram model, and 3) application of the Kriging system to produce concentration prediction surfaces and probability of exceedance maps.

2.3 GIS data preprocessing

As indicated above, the analysis in this paper is based on data collected through an extensive campaign sponsored by the USAID to assess groundwater conditions in the ULB. The data was quality checked and processed into a GIS database. The GIS facilitated further validation and quality control of the data. Using ESRI ArcGIS, the GIS data was organized into layers representing different water quality parameters for the winter and summer periods. ArcGIS Geostatistical Analyst extension was then used to carry out the three-step geostatistical analysis procedure for the groundwater nitrate concentrations. The following sections present the results of this analysis.

3 Analysis of nitrate groundwater contamination

3.1 Data exploration

Nitrate concentrations were first checked for normality. The histogram for the winter nitrate concentrations (fig 3a) indicates that the dataset is highly skewed and does not fit a normal distribution. Applying the logarithmic transformation, which is commonly used in geostatistics, to the dataset produced a bell-shaped histogram (fig 3b) that signifies that the dataset is lognormally distributed.

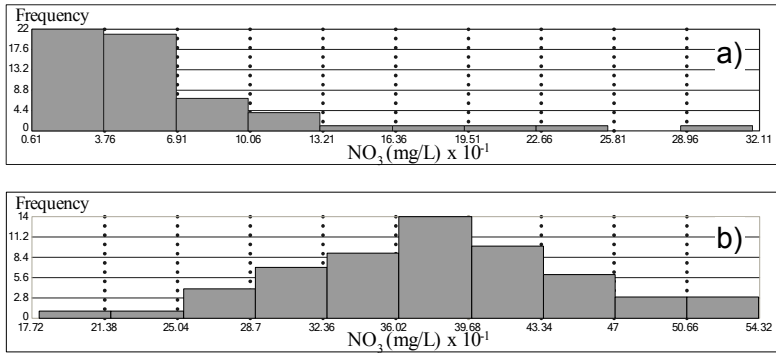


Figure 3: Histograms for the winter groundwater nitrate concentrations: a) untransformed, b) logarithmic.

Histogram analysis of the summer nitrate concentration dataset also indicated that the dataset is lognormally distributed. Based on these results, further geostatistical analysis was carried out on the log-transformed winter and summer nitrate concentration datasets.

Both datasets were also checked for the presence of global trends, but no significant trend was identified. Trend analysis is an important step for assessing the stationarity of the data. A global trend should be removed prior to carrying out structural analysis.

3.2 Structural analysis

Selecting a semivariogram model is an iterative process that involves calculating experimental semivariograms, fitting an alternative semivariogram, calculating an alternative Kriging prediction surface and carrying out a cross-validation statistical analysis to assess the performance of the prediction surface in terms of unbiasedness and estimation of uncertainty. In the cross-validation process, a data point is removed from the dataset and its value is estimated from the rest of data points. The process is repeated for all data points. Two main evaluation criteria: the Mean Standardized Prediction Error (MSPE), and the Root-Mean-Square Standardized Prediction Error (RMSSPE) were used to assess the unbiasedness and the estimation of uncertainty, respectively. The MSPE and RMSSPE are defined as follows:

$$MSPE = \frac{\sum_{i=1}^{n-1} (Z^*(x_i) - z(x_i)) / \sigma^*(x_i)}{n-1} \tag{7}$$

$$RMSSPE = \sqrt{\frac{\sum_{i=1}^{n-1} [(Z^*(x_i) - z(x_i)) / \sigma^*(x_i)]^2}{n-1}} \tag{8}$$

where $Z^*(x_i)$ and $\sigma^*(x_i)$ are the estimated value and standard error of the variable at location x_i , respectively, based on the other $n-1$ data points. The objective of the structural analysis process is to obtain an MSPE value close to zero, which indicates unbiasedness of prediction errors, and an RMSSPE value close to one, which indicates accurate estimation of prediction variability.

Based on structural analysis, a spherical model (range = 6.08 kms; sill = 0.894; nugget = 0.184) was selected for the winter semivariogram (fig. 4a). The summer semivariogram was fitted with a spherical model (range = 6.09 kms; sill = 0.759; nugget = 0.002) (fig. 4b). The spherical semivariogram model is generally defined as follows:

$$\gamma(h) = \begin{cases} \theta_p \left[\frac{3}{2} \frac{h}{\theta_r} - \frac{1}{2} \left(\frac{h}{\theta_r} \right)^3 \right] + \theta_n & \text{for } 0 \leq h \leq \theta_r \\ \theta_s & \text{for } h > \theta_r \end{cases} \tag{9}$$

where $\theta_r, \theta_s, \theta_p$ and θ_n are the range, sill, partial sill and nugget, respectively.

The semivariograms show significant spatial continuity in both winter and summer groundwater nitrate levels. This is a reflection of the high stability and mobility of nitrates in groundwater, which facilitate the migration of nitrates untransformed well beyond their source of input given the presence of highly permeable subsurface materials with adequate dissolved oxygen (Canter [4]). Subsurface layers in the ULB range from moderately permeable silty deposits along the main course of the river to highly permeable sand/gravel deposits in the north-eastern part along the Hala River and very permeable karstified limestones in the southern part of the basin (BAMAS [1]).

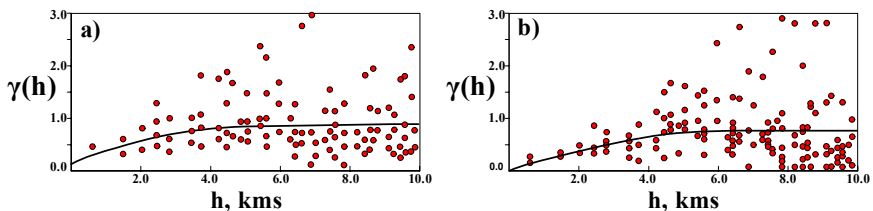


Figure 4: Semivariogram models of nitrate ground-water levels: a) winter, b) summer.



3.3 Kriging estimation of nitrate concentration maps

Applying the Kriging system to the selected semivariogram models, two prediction surfaces were calculated for the winter and summer nitrate concentrations (fig. 5). The MSPE and RMSSPE values for the winter prediction surface are -0.067 and 1.041, respectively indicating a relatively insignificant unbiasedness and a good estimation of prediction variability. The MSPE and RMSSPE for the summer prediction surface are 0.039 and 0.625 indicating an insignificant unbiasedness, but a relatively high underestimation of prediction variability, although it is within acceptable limits set by geostatistics practitioners (Cooper and Istok [5]).

The prediction surfaces show highly variable, yet very significant and persistent nitrates contamination of groundwater throughout the basin, which is mainly attributed to leaching from heavily applied fertilizers. However, several areas have excessively high nitrate levels (>100 mg/L), indicating a possible pollution by point sources.

Winter and summer contamination spatial patterns are generally similar with the exception of the northern part of the basin. Winter nitrate concentrations are generally higher than those in the summer, which reflects the lag between the application of fertilizers in the valley during the rainless season (April-August) to the time nitrates leach into the groundwater. In addition to the timing of fertilizers application, nitrates' leaching is generally affected by the soil porosity, vegetative cover, irrigation practice and rainfall intensity. In particular, the area extending east from the confluence with the Hala River shows the highest and most extensive groundwater nitrates contamination. This can be attributed to the presence of highly porous subsurface layers of sand and gravel in the area which provide poor resistance to nitrate leaching.

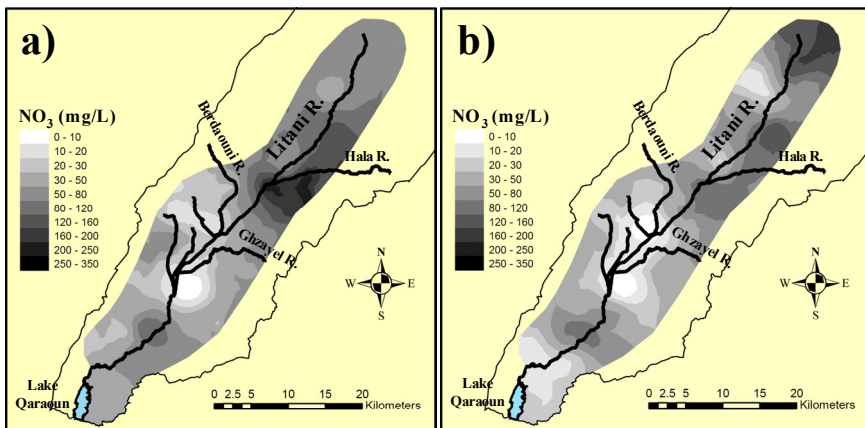


Figure 5: Nitrate concentration in groundwater in the Upper Litani Basin: a) Winter, b) Summer.

3.4 Kriging estimation of probability of exceedance maps

Groundwater quality can be characterized in terms of its adherence to specific regulatory health environmental standards. In the USA, health regulations call for drinking water supplies to contain no more than 45 mg/L of nitrates. Although the validity of this figure is currently being debated by the environmental health research community (van Grinsven *et al* [10]), it is still the *de facto* standard for drinking water in many parts of the world.

One of the advantages of using Kriging is the ability to provide assessment of the variability of estimates. Using this information, maps showing the probability of groundwater nitrate level exceeding 45 mg/L were generated for winter and summer condition (fig. 6). In winter, groundwater is most likely unsuitable for drinking in many parts of the ULB, especially in the northern and southern parts. Quality is generally acceptable in the middle section of the basin. The situation is similar, but less extensive and more localized in the summer, with the southern bordering Qaraoun Lake and the middle section maintaining the most potable groundwater conditions.

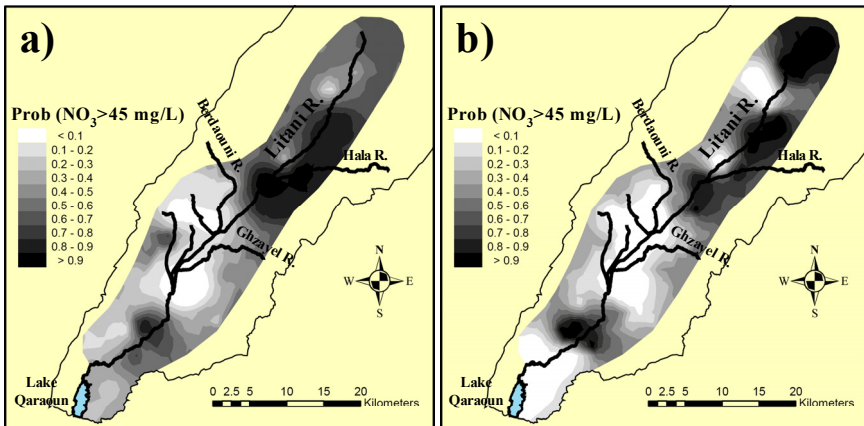


Figure 6: Probability of exceedance maps of nitrate concentrations in groundwater in the Upper Litani Basin: a) winter, b) summer.

4 Summary and conclusions

The paper presents an assessment of groundwater quality conditions in the Upper Litani Basin based on geostatistical analysis of sampled nitrate levels representing winter and summer seasons. The dataset was compiled from the outcome of an extensive sampling campaign that covered 60 sites across the basin. The paper provides a summary of geostatistics principles and methodology. Kriging, a key geostatistical tool, was used to produce georeferenced maps showing distribution of nitrate levels and probability of exceeding drinking water standards.

The results indicate significant and widespread nitrate contamination of the basin groundwater mainly attributed to leaching from excessive application of fertilizers, with a strong possibility of point source pollution. Winter nitrate levels are generally higher than summer ones signifying the lag between fertilizer application and groundwater contamination.

Groundwater is deemed unpotable for many areas in the basin, specifically in the winter and in certain sections. In contrast, groundwater in the middle section of the basin shows much lower nitrate levels, especially in the summer, which meet potable water standards.

References

- [1] Assaf, H. & Saadeh, M., Development of an integrated decision support system for water quality control in the Upper Litani Basin, Lebanon. *Proc. of the iEMSs Third Biennial Meeting: "Summit on Environmental Modelling and Software"*, eds., Voinov, A., Jakeman, A.J., Rizzoli, A.E., International Environmental Modelling and Software Society, Burlington, USA, July 2006.
- [2] BAMAS, Final Report, DAI, United States Agency for International Development, Beirut, Lebanon, 2005a.
- [3] BAMAS, Groundwater flow modelling and vulnerability, DAI, United States Agency for International Development, Beirut, Lebanon, 2005b.
- [4] Canter, L. W., *Nitrates in Groundwater*, Lewis Publishers, Boca Raton, 1997.
- [5] Cooper, R.M., Istok, J.D., Geostatistics applied to groundwater contamination. I: methodology, *Journal of Environmental Engineering, ASCE*, **114(2)**, pp. 270-286, 1988.
- [6] Friedrich, D., Chapman, D. & Beim, A., The use of Biological Material (Chapter 5). *Water Quality Assessment*, ed. D. Chapman, Chapman & Hall: Cambridge, pp. 175-242, 1996.
- [7] Journel, A.G. & Huijbregts Ch. J. *Mining geostatistics*, Academic Press, NY, 1978.
- [8] Rossi, R.E., Mulla, D. J., Journel, A.G. & Franz, E.H., Geostatistical tools for modelling and interpreting ecological spatial dependence, *Ecological Monographs*, **62(2)**, pp. 277-314, 1992.
- [9] Ward, M.H., deKok, T.M., Levallois, P., Brender, J., Gulis, G., Nolan, B.T. & VanDerslice, J., Workgroup report: drinking-water nitrate and health – recent findings and research needs. *Environmental Health Perspectives*, **113(11)**, pp. 1607–1614, 2005.
- [10] van Grinsven H.J., Ward M.H., Benjamin N. & de Kok TM, Does the evidence about health risks associated with nitrate ingestion warrant an increase of the nitrate standard for drinking water? *Environ Health*, **5(26)**, 2006.

