
Spatial assessment and redesign of a groundwater quality monitoring network using entropy theory, Gaza Strip, Palestine

Y. Mogheir · V. P. Singh · J. L. M. P. de Lima

Abstract Using entropy theory, a methodology was developed for the evaluation and redesign of groundwater quality monitoring wells in the Gaza Strip in Palestine. Essential to the methodology is the development of a Transinformation Model (TM) which yields the amount of information transfer and the dependency between the wells as a function of distance. The TM parameters, such as the minimum transinformation and the range, were employed for evaluating the network which revealed that most of the distances between wells were less than the range. It also indicated that a high percentage of redundant information existed in the network. Therefore, the network was reduced by superimposing a square pattern over the monitored area and selecting one well per square block in a stratified pattern. The methodology was tested using the chloride data collected from 1972–2000 from 417 groundwater quality monitoring wells in the Gaza Strip. The number of the groundwater quality monitoring wells in the Gaza Strip was reduced by 53%, while there was 26% redundant information based on the minimum existing distance between wells. This methodology is meant to help monitor the groundwater quality (salinity) in the Gaza Strip.

Résumé En utilisant la théorie de l'entropie on a développé une méthodologie pour évaluer et projeter un système des forage de surveillance de la qualité des eaux

souterraines dans la bande de Gaza en Palestine. L'essentiel de la méthodologie est représenté par un modèle de transformation (MT) qui rend la quantité de transfert de l'information et d'interférence entre les forages en fonction de distance. Pour évaluer le système on a utilisé les paramètres du MT comme la transformation minimale et l'écart qui ont mis en évidence que la plupart des distances entre les forages soient inférieure à l'écart. De plus le modèle a indiqué un pourcentage élevé d'information redondante dans le réseau. En conséquence le réseau a été réduit en surimposant une structure en carrés sur la zone surveillée et en sélectionnant un seul forage dans chaque block carré dans une structure stratifiée. La méthodologie a été vérifiée pour les donnés de la concentrations en chlorures, collectées dans un système de 417 forages en monitoring de la qualité des eaux dans la bande de Gaza. Cette méthodologie peut bien aider la surveillance de la qualité des eaux souterraine de la bande de Gaza.

Resumen Se desarrolló una metodología para la evaluación y rediseño de uno pozos de monitoreo de calidad en agua subterránea, en la Franja de Gaza en Palestina, mediante el uso de la teoría de la entropía. Para la metodología es esencial el desarrollo de un Modelo de Transinformación (MT), el cual entrega la cantidad de transferencia de información y la dependencia entre los pozos, como una función de la distancia. Los parámetros del MT, tales como la transinformación mínima y el rango, fueron usados para evaluar la red, lo cual reveló que muchas de las distancias entre pozos eran menores que el rango. Esto también indicó que existía un alto porcentaje de información redundante en la red. Por lo tanto, la red fue reducida mediante la sobreposición de un patrón de forma cuadrada, sobre el área monitoreada y seleccionando un pozo por cada celda cuadrada en un patrón estratificado. La metodología fue probada utilizando los datos de cloruros colectados entre 1972–2000, a partir de 417 pozos de monitoreo de calidad en agua subterránea en la franja de Gaza. Se espera que esta metodología ayude a monitorear la salinidad del agua subterránea en la Franja de Gaza.

Received: 18 April 2004 / Accepted: 5 May 2005
Published online: 16 November 2005

© Springer-Verlag 2005

Y. Mogheir
Environment Quality Authority,
Gaza, Palestine

Y. Mogheir · J. L. M. P. de Lima (✉)
Institute of Marine Research, Coimbra Interdisciplinary Centre;
Department of Civil Engineering, Faculty of Science and
Technology, Campus 2, University of Coimbra,
3030-290 Coimbra, Portugal
e-mail: plima@dec.uc.pt
Fax: +351-239797179

V. P. Singh
Department of Civil and Environmental Engineering, Louisiana
State University,
Baton Rouge, LA 70803-6405, USA

Keywords Groundwater monitoring · Entropy theory · Redundant information · Salinity · Gaza strip

Introduction

Groundwater monitoring provides valuable data which are employed for a variety of applications, such as describing, examining and documenting changes in the aquifer condition. They may permit an evaluation of recommended monitoring needs and monitoring strategies (number and location of monitoring sites, frequency of monitoring and variables to be monitored) that should be implemented to gauge the success of aquifer management, including aquifer protection from potential sources of groundwater pollution (e.g., Mogheir and Singh 2002b). The monitoring must be such that an accurate assessment of the characteristics of the aquifer (e.g., groundwater quality) can be made. This requires an a priori understanding of hydraulic and water quality regimes of the aquifer as well as consideration of all major factors (e.g., rainfall, pumping) that may affect the aquifer under present and future conditions.

Appropriate and adequate data are essential for the planning and management of aquifers. Monitoring is, therefore, closely linked to the aquifer management, since the results of monitoring may require changes or modifications in the management practice. The data collected from an aquifer monitoring network may reflect the shortage or redundancy of information. Many attempts have been made to reduce the gap between the information needs and the data collected by the monitoring network. One of the methods used for reducing this gap is the entropy theory-based method. Entropy theory quantitatively measures the information content of a series of data in terms of the reduction in uncertainty. It also defines the mutual information between two variables by means of transinformation (Mogheir et al. 2004a) which can be used to describe the spatial variation of observed variables. Transinformation can be used to evaluate the informativeness of data (insufficiency, redundancy or adequacy) with respect to its variability in space. Accordingly, the monitoring network may be expanded if there is insufficient information, or it may be reduced if there is redundant information.

Selection of the number and location of the groundwater quality monitoring wells requires the consideration of different aspects, such as monitoring objectives and the cost of monitoring (Mogheir 2003). The cost of monitoring was considered in Mogheir et al. (2003b) where the assessment of the monitoring network recommended a reduction of the network and consequently a reduction in the cost. Moreover, Mogheir (2003) proposed a multi-objective design approach which showed that the entropy-based approach can be effectively applied for monitoring different variables, such as EC and NO₃ (indicators for other pollutants).

Therefore, the main objective of this paper is to present a step-by-step methodology for assessing and redesigning the location of groundwater quality monitoring wells using the entropy theory. A case study from Gaza Strip, Palestine, illustrates an application of the methodology described. A regional (national) approach for the assessment

and redesign of the groundwater monitoring network has been adapted to the Gaza Strip to monitor the salinity in the existing aquifer.

Literature review

There is a multitude of approaches to the design of hydrologic monitoring networks. A brief review of these approaches is in order.

Geostatistical methods

Essential for the design approaches is a clear definition of the objective of monitoring. Examples of monitoring objectives (Loaiciga et al. 1992) are ambient monitoring (gathering information at a regional scale), detection monitoring (for point and non-point pollution sources), compliance monitoring (verifying the process of remediation projects at the local scale), and research monitoring (spatial and temporal design to meet specific goals). The objective of monitoring should be based on the objective functions (ultimate objectives), such as environmental protection (e.g., Hsueh and Rajagopal 1988), reduction of remediation costs (e.g., Reichard and Eves 1989), and minimizing the exposure risk or minimizing health hazard (e.g., Massmann et al. 1991). These functions should be expressed in terms of statistical parameters which can be determined through the design approach. Design criteria may include minimizing the variance of contaminant concentrations (e.g., Rouhani 1985) and minimizing the maximum absolute deviation between actual and predicted variables, such as mean or average of estimated square errors (e.g., Loaiciga et al. 1992).

The sampling program (in space and time) must be revised and updated in response to the change in information needs and in the data gathered. Therefore, the method for groundwater monitoring network design should include the objective function, spatial scale (regional or local) and hydrogeological considerations. Based on the aforementioned criteria, the geostatistical methods of groundwater quality monitoring network design are classified into hydrogeological and statistical approaches.

Hydrogeological approach

In this approach, the network design is based on empirical experience without the use of advanced geostatistical methods. The spatial network design depends on the location of the pollution sources. The approach is applied to site-specific applications (local scale), usually after application of groundwater flow and transport models (e.g., Everett 1980).

Statistical approach

The network design is based on defining the spatial structure of the variable to be monitored. This variable may vary

in both space and time. This approach is classified into simulation and variance-based, optimization, probability-based and entropy-based approaches.

Simulation approach

This approach primarily uses the simulation capability of geostatistical models. In this approach, the hydraulic conductivity is viewed as a regionalized variable. The conceptual backbone of the simulation approach is that by generating multiple synthetic fields of hydraulic conductivity, for each of which there is a corresponding contaminant field, it is possible to determine the statistical properties of mass transport in an aquifer, and thus the reliability of a monitoring network. Accordingly, the various network configurations can be entertained and examined for their adequacy in contaminant detection. The approach is tailored for problems involving contaminant monitoring at the field scale (e.g., Meyer et al. 1994).

Variance-based approach

This approach uses the estimation variance as a geometric indicator for the accuracy of spatial estimation. The approach suggests that the reduction in the estimation of variance due to additional measurements may be regarded as an improvement in the accuracy and therefore as an information gain. Thus, the objective is to minimize the estimation variance subject to some constraints. The variance-based approach is divided into global and local methods.

Global method

The method identifies the best pattern (square, triangular, or other geometric configuration) and the best density (the number of monitoring points or sampling sites). The maximum variance of estimation was used as the global design indices by Olea (1984), Yfantis et al. (1987), Christakos and Olea (1988).

Local method

This method uses a methodical search for the number and locations of sample sites that would minimize the variance or estimation error for the variable of interest. The search starts with a number of existing sampling wells to which additional wells from a pool of potential sites are added one at a time. Sampling sites continue to be added until the estimation variance can no longer be reduced. This approach is good for selecting the best group of additional sampling sites when the primary objective is to gain as much information as possible and is used for local scale problems (e.g., Rouhani and Hall 1988; Gao et al. 1996).

Optimization approach

This is a mathematical programming approach where the objective function is to minimize the estimation variance

of the groundwater quality monitoring network properties subject to specified constraints. The approach is an analytical tool for regional network design. Its shortcoming is that it does not consider the iterative nature of many typical activities of groundwater quality monitoring networks (e.g., Mahar and Datta 1997; Wang and Zheng 1997; Cunha 1999).

Probability-based approach

The selection criterion for this approach includes both accuracy (represented by the estimation variance) and the magnitude of the estimated value (such as contaminant concentration). That is done by introducing the probability of exceeding a certain level of the field variable as the criterion to be controlled in the network design. The approach seems to be suitable for regional scale problems (e.g. Angulo and Tang 1999).

Entropy-based approach

Entropy theory has been used in assessing and designing monitoring networks in different fields: water quality monitoring networks (e.g., Wu and Zidek 1992; Harmancioglu et al. 1999) and rainfall gage stations (e.g., Sounga 1976; Krastanovic and Singh 1992). Caselton and Husain (1980) reduced an existing rainfall monitoring network by selecting the combination of monitoring stations that yielded the highest joint entropy. Assuming the gamma distribution for the rainfall data, Husain (1989) used transinformation as a function of correlation to evaluate the monitoring network. Bueso et al. (1998) employed the Gaussian distribution for groundwater level data. Assuming an equivalence with the covariance, conditional entropy was used for designing the monitoring network. Thus, the fundamental criterion was to minimize the conditional entropy which amounted to minimizing the covariance in the Gaussian case. This approach was applied to reduce an existing groundwater level monitoring network. Harmancioglu and Alpaslan (1992) and Ozkul et al. (2000) quantitatively measured the network efficiency in terms of the information produced by the network. The method was applied to assess the sampling sites, temporal sampling frequencies and combined space/time surface water quality networks.

Most of the aforementioned studies that applied the entropy theory to the monitoring network design used the Gaussian distribution for the monitored variables and obtained the transinformation value as:

$$T = -0.5 \ln(1 - r^2) \quad (1)$$

where T is the transinformation, and r is the correlation coefficient (e.g., Mogheir et al. 2004a).

Using the entropy theory, Mogheir and Singh (2002a) proposed a framework for spatial design (or redesign) of groundwater monitoring networks. Marginal entropy, conditional entropy, joint entropy and transinformation were used for redesign. These were computed using the

contingency tables (An example illustrating the construction of contingency tables is presented in Mogheir et al. 2004a). A fundamental step of the design methodology was concerned with developing a spatial structure model using the entropy theory. This involved characterizing the spatial variability of the monitored variables. Transinformation was computed using contingency tables and the distance between wells. The discrete approach was compared with the analytical approach by which the transinformation was computed as a function of correlation. The analytical approach, which assumed the Gaussian distribution function, has limitations in practice, since the field groundwater data do not always fit the Gaussian distribution function and the mean does not equal zero (Mogheir et al. 2004a). To investigate the validity of the discrete approach, the methodology was also applied to two groundwater quality variables: electrical conductivity and chloride (Mogheir et al. 2003a).

The literature indicated the entropy theory can be effectively used to solve problems in water resources and environmental management, where uncertainty and information are the most significant concepts. The decisionmaking process is highly dependent on the evaluation of the available information about processes investigated. However, there has been no quantitative measure developed yet to assess uncertainty in objective terms. Thus, the two concepts are often described subjectively or indirectly in terms of other variables or parameters. Under these conditions, entropy appears to be a highly attractive concept since it directly defines information. That is basically the most significant contribution of entropy to investigations in hydrology and water resources. This feature leads to more reliable and objective evaluations within the decision-making process, especially when inferences are to be based on incomplete or inadequate information (e.g., Harmancioglu and Singh 1998; Mogheir 2003).

Groundwater monitoring in the Gaza Strip

Gaza Strip hydrogeology

The Gaza Strip, which is located in the southern corner of Palestine, is divided into five Governorates: Northern, Gaza, Middle, Khanyunes and Rafah (Fig. 1). Each Governorate has its own municipal and local administration. The water resources in the Gaza Strip are limited to the groundwater in the coastal aquifer that underlies its 365 km² area.

The coastal aquifer is the only aquifer in the Gaza Strip and is composed of Pleistocene marine sand and sandstone, interlaced with clayey layers. The maximum thickness (150 m) of water-bearing horizons occurs in the northwest along the coast and decreases gradually toward the east and southeast along the eastern border of the Gaza Strip to less than 10 m (Fig. 2). Near the coast, clay layers extend about 2–5 km inland, and divide the aquifer sequence into three or four subaquifers, depending upon the location (referred to as subaquifers A, B1, B2, and C). The base of the coastal aquifer system is formed of impervious clay shade rocks of Neogene age (Saqiyah Formation) with a range in total thickness between 500 to 1,000 m.

The depth to the groundwater level of the coastal aquifer varies between a few meters in the lowland area along the shoreline and about 70 m along the eastern border.

The coastal aquifer holds approximately 5×10^9 m³ of groundwater of varying quality. Only 1.4×10^9 m³ of this is “fresh water”, with Cl content of less than 500 mg/l (Metcalf & Eddy 2000a). This fresh groundwater typically occurs in the form of lenses that float on the top of the brackish and/or saline water. This means that approximately 70% of the aquifer water is brackish or saline and only 30% is fresh water which is found mainly in the Northern Governorate. Moreover, if the cross section in Fig. 2 is considered, the fresh water can be found in the shallow aquifer (A sub-aquifer), which is penetrated by most of the water wells (municipal and agricultural). The major source of recharge of groundwater in the aquifer is rainfall. Rainfall is sporadic across the Gaza Strip and generally varies from 400 mm/y in the north to about 200 mm/y in the south (Metcalf & Eddy 2000a).

The lateral inflow to the aquifer is estimated at between 10 to 15×10^6 m³/y (Metcalf & Eddy 2000a). Some recharge is available from a major surface water channel (Wadi Gaza). However, extensive extraction from Wadi Gaza in Israel limits this recharge to 1.5 to 2×10^6 m³/y. As a result, the total fresh water recharge at present is limited to approximately 56 to 62×10^6 m³/y (Metcalf & Eddy 2000a). Under natural conditions, groundwater flow in the Gaza Strip is towards the Mediterranean Sea, where it discharges. However, pumping for over 40 years has significantly altered natural flow patterns. In the year 2000, a large cone of depression formed in the northern and southern parts of the Gaza Strip where the groundwater levels are approximately 2 m below mean sea level.

Groundwater quality

High concentrations of salinity (Cl) and nitrate (NO₃) in the aquifer are the major water quality problem.

Chloride (salinity)

Salinity in the Gaza coastal aquifer is most often described by the Cl concentration in groundwater. Intensive exploitation of groundwater in the Gaza Strip in the past 30–40 years has disturbed the natural equilibrium between fresh water and saline water and has resulted in increasing salinity in most areas in the Gaza Strip. Depending on the location and hydrochemical processes, the rate of salinization may be gradual or sudden. Observed data show that only around 10 percent of the Gaza aquifer’s resources contain groundwater that meets the WHO water standard for Cl (250 mg/l), primarily in the north and along the dune sand in areas of the southwest (Metcalf & Eddy 2000a). The source of Cl may be (1) seawater intrusion where it extends inland in the range from 1.0 to 2.5 km in the northern and the southern parts of the Gaza Strip. These areas correspond to the largest pumping centers in the Gaza Strip, where groundwater levels are generally 1–2 m below mean sea level. (2) Lateral inflow of brackish water (Cl = 500–

Fig. 1 Location map of Gaza Strip, Palestine (Source: Metcalf & Eddy 2000a)

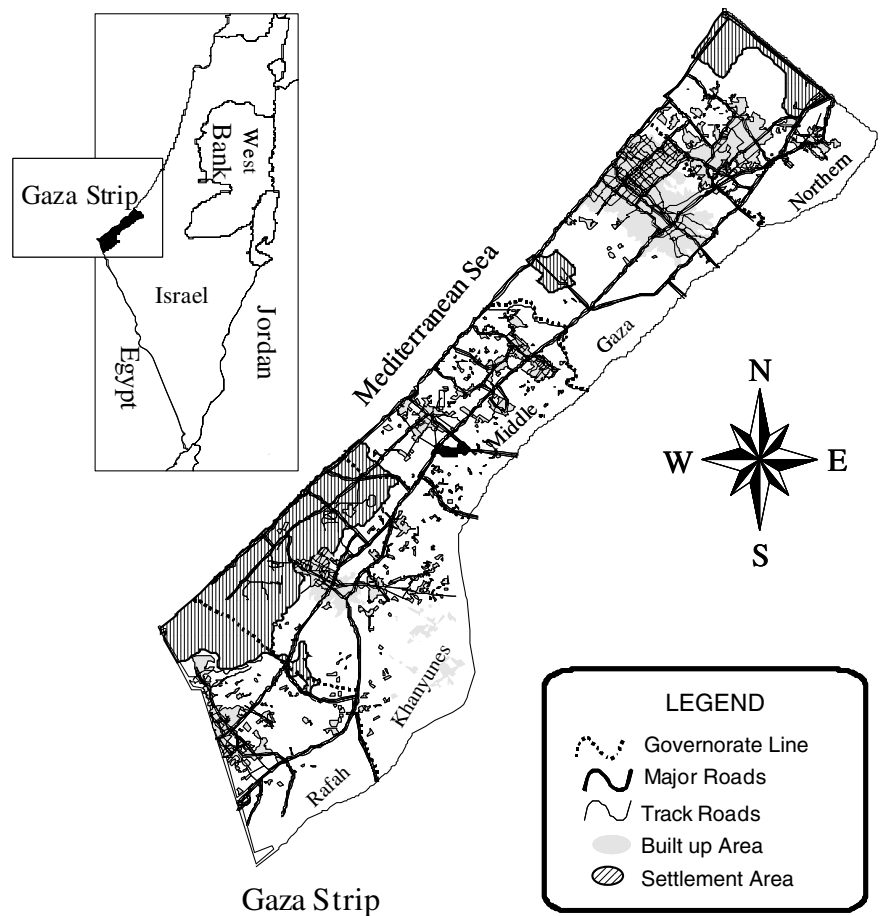
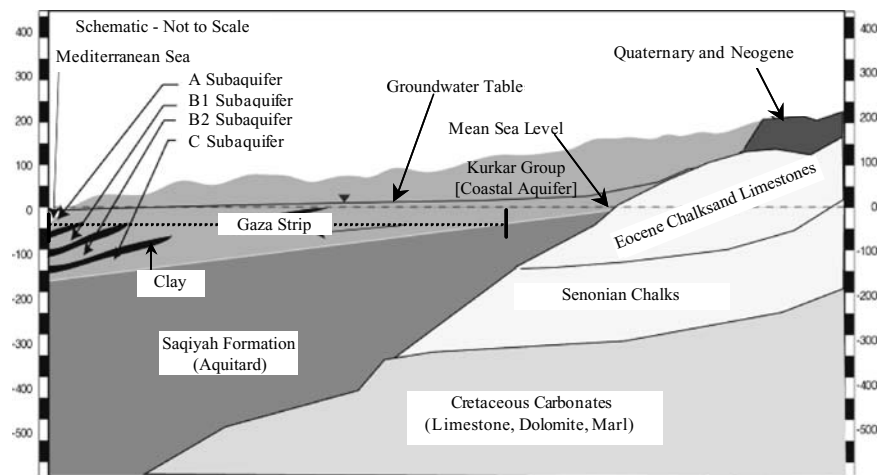


Fig. 2 Generalized hydrogeological cross section of the Gaza Strip aquifer, (Source: Metcalf & Eddy 2000a). The approximate location of Gaza Strip is indicated



2,000 mg/l) from Israel in the middle and the southern areas of the Gaza Strip. (3) Deep brines at the base of the Coastal aquifer ($Cl=4$ to 6×10^4 mg/l); overpumping results in upconing of this brine.

Nitrate (NO_3)

Most municipal wells in the Gaza Strip show nitrate levels (NO_3) in excess of the WHO drinking water standard of 50 mg/l. In the worst affected areas (urban cen-

ters), the NO_3 concentrations are increasing, in some cases rapidly, at rates of up to 10 mg/l per year. The main sources of NO_3 are fertilizers and domestic sewage effluents. The quantities of sewage that percolate to the water table on an annual basis through cesspits and septic tanks are significant, being estimated at around 12 Mm^3/y (Metcalf & Eddy 2000a). In contrast to salinity, groundwater flowing from the east has relatively low NO_3 levels.

Table 1 Summary of the existing groundwater monitoring networks in the Gaza Strip

Variables	Number of wells ^a (Approximate)	Sampling frequency	Data record since (year) ^g
Groundwater level	133 ^d	Monthly	1970
Chloride and EC	1,128 ^b 366 ^c	Twice per year	1970
Nitrate	1,128 ^b 366 ^c	Twice per year	1993
Major ions ^e	97 ^f (Mostly municipal)	Twice per year	1976

Notes:

^a95% of monitoring wells are shallow, penetrating less than 30 m into the aquifer

^bPer Ministry of Agriculture, 1,128 wells were originally sampled

^cPer Ministry of Agriculture, 366 wells were sampled since 1999

^dPer Palestinian Water Authority (PWA), 2000

^eTDS, Cl, NO₃, pH, Ca, Mg, K, F, SO₄, hardness and alkalinity

^f97 wells exist today. In 1999, only 71 were sampled

^gTime series for most data not complete; some years are missing

Existing groundwater monitoring network

In the Gaza Strip, the groundwater monitoring networks are divided according to: (1) groundwater level, (2) groundwater quality and (3) municipal wells. The groundwater level network measures the water level on a monthly base, using approximately 130 agricultural wells. The groundwater quality network measures three variables: EC, Cl and NO₃. These measurements are made twice a year using approximately 400 agricultural wells (e.g., Mogheir and Singh 2002b). The third type of monitoring network is the municipal wells which measure: EC, total dissolved solids (TDS), calcium (Ca), magnesium (Mg), sodium (Na), potassium (K), Cl, NO₃, sulphate (SO₄), alkalinity and hardness. For this purpose, approximately 100 municipal wells are used. The measurements are made twice a year in the summer and winter. The three networks are used to collect information that describes the status of the groundwater quality and quantity of the Gaza Strip Aquifer. Table 1 provides a summary of these networks.

Review of the existing groundwater quality monitoring network indicates that the majority of monitoring has been carried out in wells screened in the shallow portion of the aquifer (at the water table). Recently, groundwater wells were surveyed by the Palestinian Water Authority (PWA) and it was then found that the total number of active wells in the Gaza Strip could total more than 3,800 wells (Metcalf & Eddy 2000b). Most of the wells, used for agricultural purposes, are shallow and are typically screened 10–20 m below the water table. As mentioned in Table 1, approximately 1,128 wells have been dedicated to monitoring the water quality status in the aquifer, mainly the salinity, since 1970. The saltwater intrusion and lateral inflow can be considered as regional phenomena and initially require a regional approach (the approach of this study). That is essential for determining and predicting the future deterioration and assessing the impact of management actions on the aquifer. Monitoring of saline water upconing requires a different approach, as upconing is primarily a local process beneath wells. The principle of monitoring upconing is the same as that for seawater intrusion and lateral flow; however, the monitoring wells have to be installed within major pumping centers and near pumping wells.

All historical Cl monitoring data from 1,128 agricultural wells were checked and processed. Significant data quality checks and corrections have been made. In these wells, the starting dates for collecting the Cl data were not the same. Moreover, some of these wells have different terminal reading dates. In addition, within the period of measurements, there are missing data. After applying the quality control and assurance to the Cl data, 417 wells are considered as an existing groundwater quality monitoring network in the Gaza Strip to measure the regional salinity occurrence in the aquifer. All the wells considered in the analyses are shallow. The very few deep wells are not included in the approach of this study, since they might represent hydrogeological characteristics of a different subaquifer.

Purpose of the current work

This paper presents further progress of the earlier work by Mogheir et al. (2004b) and Mogheir et al. (2003a) towards the use of entropy theory in the assessment and redesign of groundwater quality monitoring networks. The methodology involves two parts: (1) assessing the monitoring network, and (2) redesigning the monitoring network. The assessment of the monitoring network was based on constructing the T-Model of chloride (Cl) data collected from 417 monitoring wells distributed over 365 km² of the Gaza Strip area. The T-model parameters, mainly the minimum transformation (T_{min}) and the range (L), are employed to assess the net redundant information, which exists in the monitoring network. This reveals whether the monitoring network requires reduction or extension. The second part of the methodology employs L and the redundant information percentage as design criteria for redesigning the monitoring network. The monitoring network is redesigned by a superposition approach in which a square grid is superimposed over the monitored area and selecting one well per square block in a stratified way. Different design scenarios are illustrated using different square grids. One of these scenarios is the best that was found or the recommended scenario in which the square grid is obtained from the range (L) of the T-Model.

Development of methodology

The location and number of the existing groundwater quality monitoring wells in the Gaza Strip are assessed on the basis of the redundancy of information. The wells are re-designed according to the minimum allowable redundant information between the wells.

Spatial dependency assessment

The assessment of the location of the groundwater quality monitoring wells was made for each variable (e.g., Cl) observed in a number of monitoring wells, because Cl was considered to be an indicator of the salinity in the aquifer. The approach was to assess the reduction in the mutual information between the wells due to the presence of statistical dependency between them. Such a reduction is equivalent to the reduction in the transinformation in the series of the same variable at different locations. Thus, the objective in spatial orientation is to minimize the redundant information by appropriately choosing the number and the location of monitoring wells. That can be achieved by characterizing the spatial variability of the variable at different locations. This spatial variability is described by means of transinformation as a function of distance between wells (Mogheir et al. 2003a). The following procedure was applied to assess the spatial location of the groundwater quality monitoring network:

Step 1. It was assumed that there are Z monitoring wells that are used to monitor the groundwater quality in the aquifer (regional scale). The variable (Cl) is represented by x_{zi} , where z ($z=1, \dots, Z$) denotes the well identification number and i the time-point along the sample of size n ($i=1, 2, \dots, n$). Also, X_{hj} was defined as the value of this variable observed at d (m) away from x_{zi} , where $h \neq z$ and $j=1, 2, \dots, n$, defined in the same probability space, the discrete probabilities of occurrence of both wells are and $p(x_{zi})$ and $p(x_{hj})$, respectively, and $p(x_{zi}, x_{hj})$ is the joint probability of x_{zi} and x_{hj} . Transinformation (T) is determined as:

$$T(x_{zi}, x_{hj}) = - \sum_{i=1}^n \sum_{j=1}^n p(x_{zi}, x_{hj}) \ln \left[\frac{p(x_{zi}, x_{hj})}{p(x_{zi}) p(x_{hj})} \right] \quad (2)$$

If $h=z$, then $p(x_z)=p(x_h)$ and transinformation equals the marginal entropy of a certain well which can be determined as:

$$H(x_z) = H(x_h) = - \sum_{i=1}^n p(x_{zi}) \ln p(x_{zi}) \quad (3)$$

The joint probability $p(x_{zi}, x_{hj})$ is obtained using contingency tables (e.g., Gokhale and Kullback 1978; Mogheir and Singh 2002a), which records the frequency of the values that fall into each possible combination of two cate-

Table 2 Two-dimensional contingency table (frequency)

		x_h					
		1	2	3	...	u	Total:
x_z	1	f_{11}	f_{12}	f_{13}	...	F_{1u}	f_1
	2	f_{21}	f_{22}	f_{23}	...	F_{2u}	f_2
	3	f_{31}	f_{32}	f_{33}	...	F_{3u}	f_3
	⋮	⋮	⋮	⋮	⋮	⋮	⋮
	v	f_{v1}	f_{v2}	f_{v3}	...	f_{vu}	f_v
	Total:	$f_{.1}$	$f_{.2}$	$f_{.3}$...	f_u	n

gories. To construct a two-dimensional contingency table (Table 2):

- Let i in the sample of size n ($i=1, 2, \dots, n$) in x_z have values consisting of v categories (class intervals), while j in the sample of size n ($j=1, 2, \dots, n$) in x_h is assumed to have u categories (class intervals).
- The cell density or the joint frequency for (v, u) is denoted by f_{vu} , $v=1, 2, \dots, V$; $u=1, 2, \dots, U$, where the first subscript refers to the row and the second subscript to the column. The cell density is the count of measurements with the corresponding u class interval of the variable in x_h and v class interval of the variable in x_z .
- Marginal frequencies are denoted by f_v and f_u for the row and the column values of the variable, respectively. f_v is the summation of the cell density for each v categories of the variable in x_z .
- The joint probability $p(x_{zv}, x_{hu})$ is computed by:

$$p(x_{zv}, x_{hu}) = \frac{f_{vu}}{n} \quad (4)$$

where n is the sample size of the time series of the variable in each well. Mogheir et al. (2004a) indicated the construction of the contingency tables with reference to an example of the Gaza Strip groundwater quality data.

Step 2. The spatial dependency was measured by the T-Model. This model was determined by computing transinformation and the distance between wells. The geometrical distance (d) between two wells was calculated as

$$d = \sqrt{(A_1 - A_2)^2 + (B_1 - B_2)^2} \quad (5)$$

where A_1 and B_1 are the coordinates of a well which has outcome x_z ; A_2 and B_2 are the coordinates of a well that has outcome x_h ; and d is the distance between the two wells.

Step 3. The exponential decay curve was fitted to the T-Model using the least square fitting procedure with the GRAPHPAD PRISM statistical software (Motulsky 1999). The coefficient of determination was used to quantify the goodness of fit between the exponential decay curve and discrete models. The coefficient of determination (R^2) was computed as (e.g., Motulsky 1999):

$$R^2 = 1.0 - \frac{SS_{\text{reg}}}{SS_{\text{tot}}} \quad (6)$$

where SS_{reg} is the sum of the squares of the residuals between the discrete model and the best fit exponential decay curve, and SS_{tot} is the sum of the squares of the residuals between the discrete model and the horizontal line through the mean.

Step 4. The exponential decay of the T-Model (TM_{ED}) was represented as:

$$T(d) = (T_0 - T_{min})e^{(-Kd)} + T_{min} \quad (7)$$

where T_0 is the initial value of transinformation; K is the transinformation decay rate; T_{min} is the minimum transinformation value (T_{min}); and d is the distance between wells. The distance at which the transinformation model approaches a constant minimum value is called the range (L). An accurate estimation of the range L plays an important role for the assessment of the spatial location of the monitoring wells. For an exponential decay curve, $T(d)$ approaches asymptotically T_{min} , therefore, the L value is estimated when $T(d)$ is very close to T_{min} . Assume that at $d=L$, $T(L)=T_{min}+\varepsilon$, where ε (error term) is a constant number and its unit is Nats (natural units). If $T(L)$ is substituted in Eq. (7) by $T_{min}+\varepsilon$, Eq. (7) can be written as

$$\begin{aligned} T_{min} + \varepsilon &= (T_0 - T_{min})e^{(-KL)} + T_{min} \\ \varepsilon &= (T_0 - T_{min})e^{(-KL)} \end{aligned} \quad (8)$$

where ε can be considered the threshold which enables the estimation of L .

By taking the logarithm of both sides of Eq. (8) the following equation was obtained:

$$\ln(\varepsilon) = \ln(T_0 - T_{min}) - KL \quad (9)$$

Then the value of L can be computed as,

$$L = \frac{\ln(T_0 - T_{min}) - \ln(\varepsilon)}{K} \quad (10)$$

Step 5. The wells were assessed using the percentage of Redundant Information (% RI) between wells. At a distance $d(m)$ between wells, the % RI was computed as:

$$\%RI = \frac{T(d)}{T_0} \times 100 \quad (11)$$

For more accurate computation, one should compute the net transinformation, since the T-Model steadily decreases to a minimum value (T_{min}) which differs from zero. Practically, that is not accurate, since the minimum transinformation should equal zero. Therefore, the Net T is computed as:

$$T(d)^{NET} = T(d) - T_{min} \quad (12)$$

Then, the percentage of the Net Redundant Information (% NRI) is computed as:

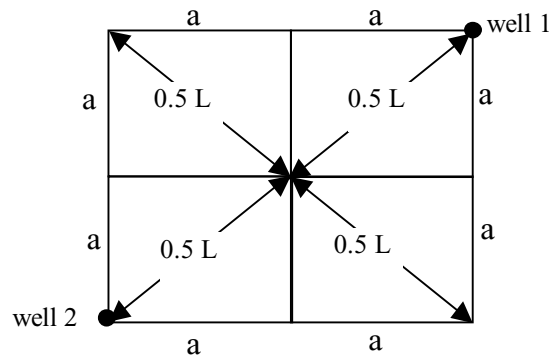


Fig. 3 Square grid to be superimposed over the existing monitoring wells. a is the grid size and L is the recommended distance between wells

$$\% NRI = \frac{T(d)^{NET}}{T_0 - T_{min}} \times 100 \quad (13)$$

The maximum NRI which can be available in the network is computed according to the minimum existing distance between two wells. The NRI value indicates further steps required for redesigning the location of the monitoring network, either by reducing or expanding the monitoring network. Note that when $d=L$, the Net T equals zero, since after the L distance one starts to lose information and the redundant information approaches zero.

Redesign of spatial location of monitoring wells

After determining L , several approaches can be used to select the number of wells from an existing monitoring network that satisfies the design criterion (distance between wells $\leq L$). One approach is the superimposition approach (e.g., Sophocleous 1983; another case by Olea 1999). This approach can be used as follows:

- From L , a regular square grid is specified (Fig. 3). Assume that the size of the square in the grid is (a) and the diagonal of the square grid is ($0.5 L$), where L is the maximum distance between two wells (well 1 and well 2 in Fig. 3). Then, the recommended grid size is obtained as:

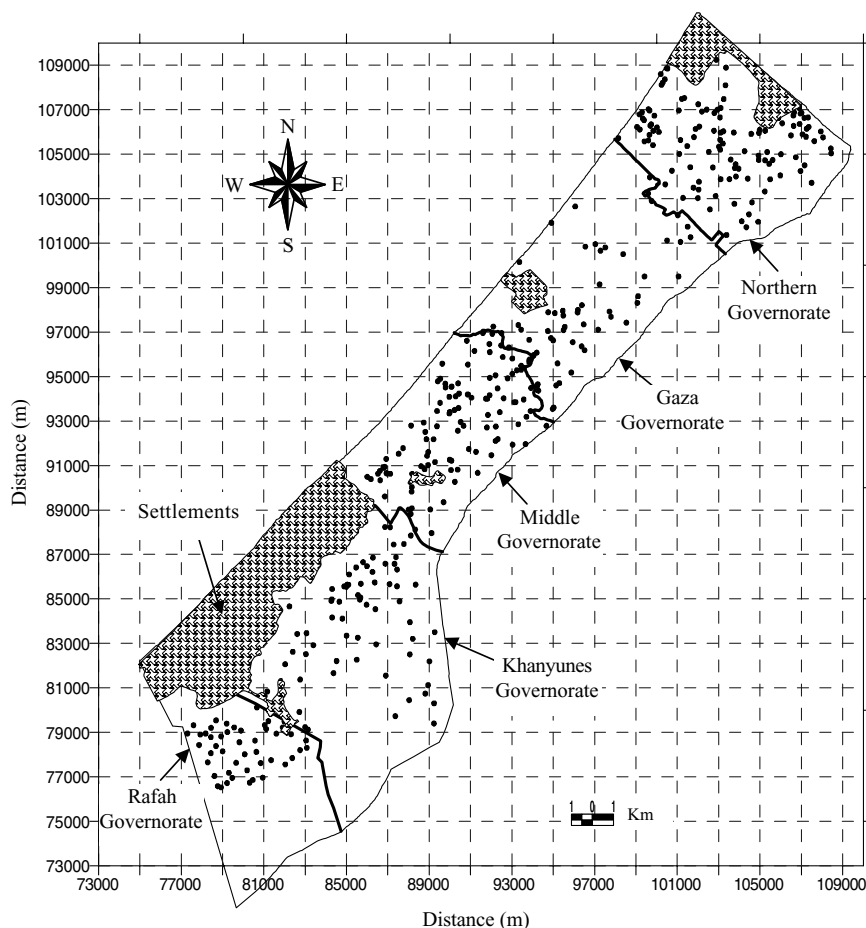
$$a = 0.5(L) \cos 45 = 0.354 L \quad (14)$$

- After determining the square grid size (a), the number of wells is reduced by superimposing a square $a \times a$ m² over the entire monitored area and selecting one well per square block in a stratified way. The final square grid reduces the number of existing wells as:

$$\%Reductioninwells = [1 - (\text{numberofselectedwells} / \text{numberofexistingwells})] \times 100 \quad (15)$$

- Through the superimposition approach, different scenarios were evaluated by specifying different grid sizes. As a result, different design options were identified, which

Fig. 4 The location of the existing groundwater quality monitoring wells (black dots) in the Gaza Strip area



gave different values of % of NRI and % of reduction in wells. As part of the scenarios, the recommended design option was obtained based on L (Eq. (14)).

- There are areas where the distance is greater than L and this area needs additional monitoring wells (expansion of the network). The additional wells may be located in the center of the square grid in the empty area.

Application of method

The above methodology was applied to the Gaza Strip where the existing groundwater quality monitoring network comprises 417 wells. Figure 4 shows the location of these wells which are monitored for salinity by the Ministry of Agriculture. The analysis was applied to CI for the years 1972–2000 on a bimonthly basis. The analysis was applied to each Governorate of the Gaza Strip: Northern, Gaza, Middle, Khanyunes, and Rafah (see Figs. 1 and 4).

Analyzing spatial locations

The T-Model for CI in each Governorate was obtained by computing the T values using Eq. (2) and the distance between wells. The exponential decay curve was fitted to the T-Model data using the least square technique. The T-

Table 3 Fitting transinformation data with exponential decay curve for CI in each Governorate in the Gaza Strip area

Governorate	Fitting equation ^a	R^2
Northern	$T(d) = 0.9860 e^{(-0.0034 d)} + 0.4583$	0.55
Gaza	$T(d) = 1.0900 e^{(-0.0034 d)} + 0.4768$	0.57
Middle	$T(d) = 0.9600 e^{(-0.0034 d)} + 0.5836$	0.44
Khanyunes	$T(d) = 1.0720 e^{(-0.0034 d)} + 0.3997$	0.65
Rafah	$T(d) = 0.9953 e^{(-0.0034 d)} + 0.4827$	0.46

^a $T(d)$ in Nats; d in m

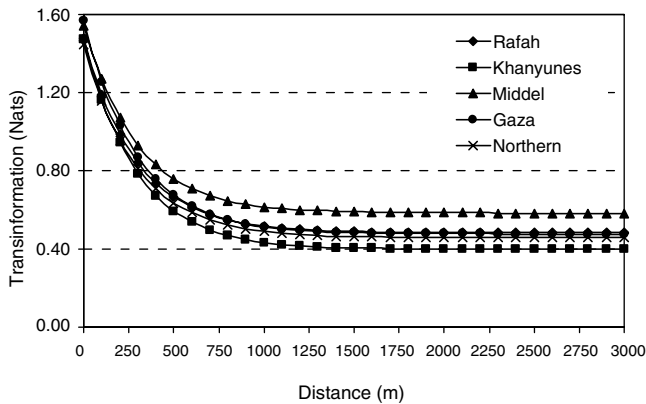
Model of CI in each Governorate is presented in Fig. 5 and Table 3.

Determining L (the range)

An accurate determination of L helps signify the degree of dependency between the wells (spatial dependency). From Fig. 5 and Table 3, the T-Model parameters, mainly T_0 , T_{\min} and K can be obtained. Equation (10) is used to determine the value of L . In this Equation ϵ is considered small and equals 0.0001 Nats. Under these conditions, the L value of each T-Model in each Governorate was selected at the point where the T-Model asymptotically reached the distance axis. Table 4 shows the parameters (T_0 , T_{\min} , K and L)

Table 4 TM_{ED} parameters for CI in each Governorate

Governorate	T_0 (Nats)	T_{min} (Nats)	K (1/m)	L (m)
Northern	1.4443	0.4583	0.0034	2705
Gaza	1.5668	0.4768	0.0034	2734
Middle	1.5436	0.5836	0.0034	2697
Khanyunes	1.4705	0.3997	0.0034	2729
Rafah	1.4780	0.4827	0.0034	2708

**Fig. 5** Transinformation Model for the CI variable in each Governorate in the Gaza Strip (Rafah, Khanyunes, Middle, Gaza and Northern)

of the exponential decay of the T-Model (TM_{ED}) in each Governorate.

Assessment of spatial locations

The minimum existing distance between wells in each Governorate was found to be less than the L value. For example, in the Northern Governorate, the minimum distance was 48.82 m. Then, the maximum redundant information (net), which was computed using Eq. (13), equalled 84.59%. Table 5 shows the evaluation of the existing monitoring network by identifying the minimum distance, the maximum Net Redundant Information (% NRI) and the number of existing monitoring wells.

Table 5 shows that the minimum distance in the Gaza Strip Governorates ranges between 24.41 and 166.85 m and the maximum NRI percentage ranges between 56.71 and 92.04%. This indicates that the location of the existing monitoring wells in the Gaza Strip Governorates needs redesigning by reducing the number of wells. The reduction was based on the criterion that a minimum redundant

Table 5 Evaluation of the existing monitoring network in each Governorate

Governorate	Minimum distance (m)	Maximum NRI (%)	Number of wells
Northern	48.82	84.59	124
Gaza	78.24	69.17	50
Middle	24.41	92.04	121
Khanyunes	54.45	83.10	71
Rafah	166.85	56.71	51

information should be allowed between wells. The decisionmaker may decide on the amount of redundant information to be followed. To overcome this issue, the designer may provide several scenarios from which one can select the number and location of wells that give the minimum redundant information in the network.

Redesign of spatial locations

Based on the assessment of the spatial locations, it was found that the spatial location of existing monitoring wells was inadequate and significant redundant information existed. Therefore, a systematic network should be determined and should offer less redundant information. The existing monitoring wells should be reduced to allow a minimum NRI percentage. That was achieved by considering the T-model parameters, specifically T_{min} and L . The redesign criterion was to select the distance between wells, which gave the minimum NRI.

The superimposition approach was used to redesign the monitoring wells by specifying different scenarios based on specified grid sizes. For each Governorate, four scenarios were evaluated using different grid sizes (500 m, 1,000 m, 1,500 m and (a) m). In each scenario, a regular square grid was superimposed over the entire monitored area and one well per square block was selected in a stratified block pattern. Different scenarios gave different design outputs: % of NRI, using Eq. (13), and % of reductions in wells, using Eq. (15), and for example, in the first scenario and the Northern Governorate, a 500 m \times 500 m square grid was superimposed over the monitored area and one well per square block in a stratified pattern was selected. In this scenario, the existing network was reduced by 28% and 49.5% of redundant information still available between wells, since the minimum distance between wells at this scenario was 210 m. The four scenarios for each Governorate are presented in Table 6.

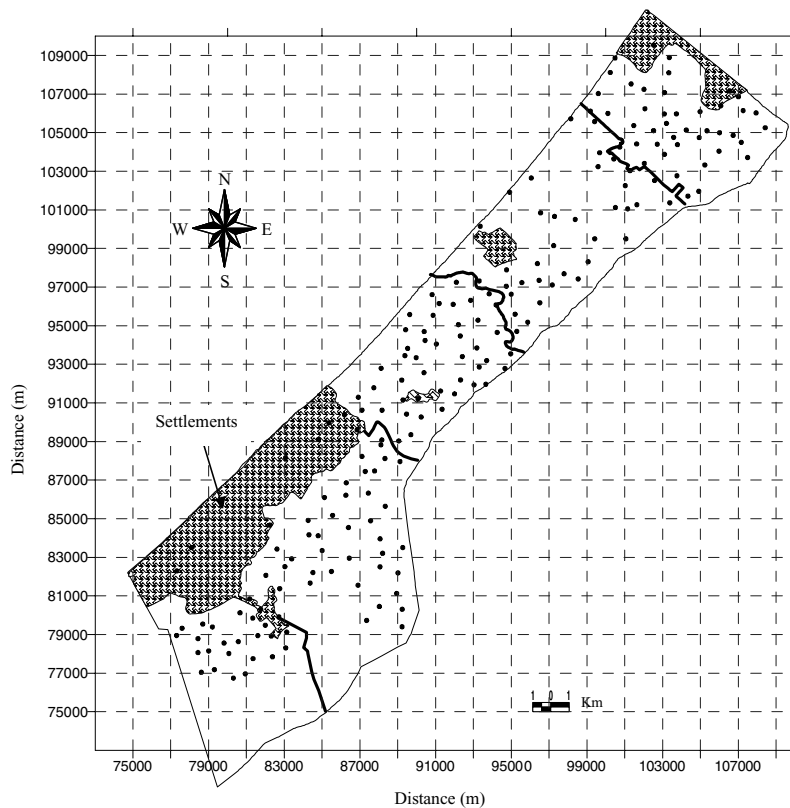
The fourth scenario is the recommended scenario. This was attained by selecting the distance between wells equal to L . For example, in the Northern Governorate, from L which was estimated as 2,705 m, a regular square grid 958 m \times 958 m was specified using Eq. (14). Then, the network in the Governorate was reduced by superimposing the square pattern over the monitored area and selecting one well per square block in a stratified way. The existing network was reduced by 60%, using Eq. (15). The NRI was 23%, which was not reduced completely, since the minimum distance between wells is 428 m. There are areas where the distance is greater than L and this area needs additional monitoring wells. These wells can be selected from those agricultural wells that have not been used for monitoring before or those agricultural wells which are used for monitoring but have some missing data. The recommended location for the monitoring wells in the Gaza Strip based on the fourth scenario is illustrated in Fig. 6.

The existing sampling density is computed as 1.90 wells/km² considering the monitored area in the Gaza Strip as 219 km². This area is about 60% of the Gaza Strip area, since the rest of the area is either

Table 6 Redesign scenarios applied to the existing monitoring network in the Gaza Strip Governorates

	Scenarios	a (Grid size) (m)	Max. distance (m)	Min. existing distance (Existing)	NRI (%)	Reduction in no. of wells (%)
Northern Governorate	Scenario 1	500	1,414	210	50	38
	Scenario 2	1,000	2,828	540	16	63
	Scenario 3	1,500	4,242	814	7	78
	Recommended Scenario	958	2,705 = L_{max}	428	23	60
Gaza Governorate	Scenario 1	500	1,414	222	42	20
	Scenario 2	1,000	2,828	385	24	40
	Scenario 3	1,500	4,242	701	8	52
	Recommended Scenario	968	2,734 = L_{max}	491	17	36
Middle Governorate	Scenario 1	500	1,414	179	54	29
	Scenario 2	1,000	2,828	434	23	61
	Scenario 3	1,500	4,242	567	15	77
	Recommended Scenario	955	2,697 = L_{max}	397	26	60
Khanyunes Governorate	Scenario 1	500	1,414	213	48	13
	Scenario 2	1,000	2,828	512	17	38
	Scenario 3	1,500	4,242	712	9	56
	Recommended Scenario	966	2,729 = L_{max}	493	19	38
Rafah Governorate	Scenario 1	500	1,414	334	32	20
	Scenario 2	1,000	2,828	509	18	59
	Scenario 3	1,500	4,242	819	6	76
	Recommended Scenario	959	2,708 = L_{max}	473	20	53

Fig. 6 The location of the recommended network of groundwater quality monitoring wells (*black dots*) in the Gaza Strip area. The number of wells is 197 and the existing wells were shown in Fig. 4



under Israeli settlements or is a military zone area. From Fig. 6 and Table 6, it can be summarized that the recommended number of groundwater quality monitoring wells is 197; and the recommended sampling density is 0.90 wells/km².

Conclusions

A methodology was developed using the entropy theory for assessing and redesigning a regional groundwater water quality monitoring network. The methodology was

applied to groundwater quality data for chloride (Cl) collected from monitoring wells in the Gaza Strip. It quantifies the spatial variability of groundwater quality variables (Cl as an example) by transinformation as a function of distance (T-Model). The T-Model is expressed as an exponential decay curve which has four important parameters: the initial transinformation, the minimum transinformation, the transinformation decay rate, and the distance at which the T-Model reaches its minimum value. An accurate estimation of these parameters is essential for an accurate assessment and redesign of the location of the monitoring wells. The minimum transinformation value and the distance at which the T Model reaches its minimum value are the main parameters by which the redundant information percentage can be computed. Accordingly, the existing network of monitoring wells can be reduced or extended. In the Gaza Strip case, the maximum redundant information percentage is 92% from which it can be concluded that the network should be reduced to attain the minimum redundant information. For all Gaza Strip Governorates, this reduction can be achieved for different scenarios or selecting one well per 0.92 km² (stratified pattern). For the Gaza Strip area the methodology introduces a uniform distribution of the wells and flexibility in maximizing the use of one existing well within a 960 × 960 m area.

The methodology used in this study is general and can, in principle, be applied to any type of network analysis, in which the objective of monitoring indicates a regional scope. Examples of these networks are groundwater level and rainfall monitoring networks. The efficiency and the accuracy of the methodology of this article might be tested and compared with the geostatistical methods outlined in the literature review section of this article. The application of these methods may give different numbers and locations of wells. That will be handled by further and future research.

Acknowledgments The authors gratefully thank the Foundation for Science and Technology of the Portuguese Ministry of Science and Higher Education for sponsoring the fellowship (Reference: SFRH/BD/6089/2001). The fellowship was provided for the first author's PhD programme entitled "Quantification of Information for Groundwater Quality Networks." The programme was undertaken at the Department of Civil Engineering, Faculty of Science and Technology, University of Coimbra, Coimbra, Portugal.

References

- Angulo M, Tang W (1999) Optimal groundwater detection monitoring system design under uncertainty. *J Geotech Geoenviron Eng*, ASCE 125(6):510–517
- Bueso MC, Angulo JM, Alonso FJ (1998) A state-space model approach to optimum spatial sampling design based on entropy. *Environ Ecol Stat* 5:29–44
- Caselton WF, Husain T (1980) Hydrologic network: information transmission. *J Water Resour Plann Manag Div*, ASCE 106(WR2):503–529
- Christakos G, Olea RA (1988) A multiple objective optimal exploration strategy. *Math Comput Model* 11(4):413–418
- Cunha MC (1999) On solving aquifer management problems with simulated annealing algorithms. *Water Resour Manag* 13(3):153–169
- Everett LG (1980) Groundwater monitoring. General Electric Company, Schenectady, New York
- Gao H, Wang J, Zhao P (1996) The update Kriging variance and optimal sampling design. *Math Geol* 28(3):295–313
- Gokhale D, Kullback S (1978) The information in Contingency Tables. Marcel Dekker Inc., New York
- Harmancioglu NB, Alpaslan N (1992) Water quality monitoring network design. *Water Resour Bull* 28(1):179–192
- Harmancioglu NB, Fistikoglu O, Ozkul SD, Singh VP, Alpaslan N (1999) Water Quality Monitoring Network Design. Kluwer Academic Publishers, Boston
- Harmancioglu NB, Singh VP (1998) Entropy in environmental and water resources. In: Herschy RW, Fairbridge RW (eds) *Encyclopaedia of hydrology and water resources*. Dordrecht, Kluwer
- Hsueh YW, Rajagopal R (1988) Modelling groundwater quality decisions. *Groundwater Monitor Rev* 8(4):121–134
- Husain T (1989) Hydrologic uncertainty measure and network design. *Water Resour Bull* 25(3):527–534
- Krstanovic PF, Singh VP (1992) Evaluation of rainfall networks using entropy: II. applications. *Water Resour Manag* 6:295–314
- Loaiciga H, Charbeneau JR, Everett GL, Fogg EG, Hobbs FB, Rouhani S (1992) Review of groundwater quality monitoring network design. *J Hydr Eng*, ASCE 118(1):11–37
- Mahar PS, Datta B (1997) Optimal monitoring and groundwater pollution source identification. *J Water Resour Plann Manag*, ASCE 23(4):199–207
- Massmann J, Freeze RA, Smith L, Sperling T, James B (1991) Hydrogeological decision analyses: 2. Applications to groundwater contamination. *Ground Water* 29(4):536–548
- Metcalf & Eddy (2000a) Coastal Aquifer Management Program, Final Report: Modelling of Gaza Strip Aquifer. The program is funded by US Agency for International Development (USAID) and owned by the Palestinian Water Authority (PWA). Gaza, Palestine
- Metcalf & Eddy (2000b) Coastal Aquifer Management Program, Aquifer Monitoring Plan. The program is funded by US Agency for International Development (USAID) and owned by the Palestinian Water Authority (PWA). Gaza, Palestine
- Meyer PD, Valchoni AJ, Eheart JW (1994) Monitoring network design to provide initial detection of groundwater contamination. *Water Resour Res* 30(9):2647–2659
- Mogheir Y (2003) Assessment and Redesign of Groundwater Quality Monitoring Networks Using the Entropy Theory—Gaza Strip Case Study. Ph.D. Thesis, University of Coimbra, Coimbra, Portugal, 319 pp
- Mogheir Y, de Lima JLMP, Singh VP (2003a) Applying the entropy theory for describing the spatial structure of groundwater regionalized variables (EC and Chloride). In: Neves MV, Neves ACV (eds) *Proc International Symposium "Environment 2010: Situation and Perspectives for the European Union"*, G 01
- Mogheir Y, Singh VP, de Lima JLMP (2003b) Redesigning the Gaza Strip groundwater quality monitoring network using entropy. In Singh VP, Yadava RY (eds), *Groundwater pollution*, Allied Publishers Pvt. Limited, New Delhi, India 5:315–331
- Mogheir Y, de Lima JLMP, Singh VP (2004a) Characterizing the spatial variability of groundwater quality using the entropy theory: I. Synthetic data. *Hydrol Process*, 18:2165–2179
- Mogheir Y, de Lima JLMP, Singh VP (2004b) Characterizing the spatial variability of groundwater quality using the entropy theory: II. Case study from Gaza Strip. *Hydrol Process*, 18:2579–2590
- Mogheir Y, Singh VP (2002a) Application of information theory to groundwater quality monitoring networks. *Water Resour Manag* 16(1):37–49
- Mogheir Y, Singh VP (2002b) Specification of information needs for groundwater resources management and planning in developing country: Gaza Strip case study. In: Sherif MM, Singh VP, Al-Rashed M (eds) *Proceedings Water Resources Management in Arid Regions Conference (2)* 3–20

- Motulsky HJ (1999) Analysing Data with GraphPad Prism. GraphPad Software. San Diego, <http://www.graphpad.com>
- Olea RA (1984) Sampling design optimisation for spatial functions. *Math Geol* 16(4):369–393
- Olea RA (1999) Geostatistics for engineers and earth scientists. Kluwer Academic Publishers, Boston
- Ozkul S, Harmancioglu NB, Singh VP (2000) Entropy-based assessment of water quality monitoring networks. *J Hydrol Eng, ASCE* 5:90–100
- Reichard EG, Eves JS (1989) Assessing the value of hydrogeologic information for risk-based action decision. *Water Resour Res* 25(7):1451–1460
- Rouhani S (1985) Variance reduction analysis. *Water Resour Res* 21:837–846
- Rouhani S, Hall TJ (1988) Geostatistical schemes for groundwater sampling. *J Hydrol* 103:85–102
- Sophocleous M (1983) Groundwater observation network design for the Kansas groundwater management districts, USA. *J Hydrol* 61:371–389
- Sounga JO (1976) Entropy principle applied to the rainfall-runoff process. *J Hydrol* 30:81–94
- Wang M, Zheng C (1997) Optimal remediation policy selection under general conditions. *Ground Water* 35(5):757–764
- Wu S, Zidek VJ (1992) An entropy-based analysis of data from selected NADP/NTN network sites for 1983–1986. *Atmos Environ* 26A(11):2089–2103
- Yfantis EA, Flatman GT, Behar JV (1987) Efficiency of Kriging estimation for square, triangular and hexagonal grids. *Math Geol* 19(3):183–205