Spatio-temporal trends and change factors of groundwater quality in an arid area with peat rich aquifers: Emergence of water environmental problems in Tanta District, Egypt

Alaa A. Masoud [a, *], Katsuaki Koike [b], Hamdy A. Mashaly [c], Fibi Gergis [a]

[a] Geology Department, Faculty of Science, Tanta University, 31527 Tanta, Egypt
[b] Graduate School of Engineering, Kyoto University, Kyoto 615-8540, Japan
[c] Holding Company for Drinking Water and Wastewater (HCDWW) Headquarter in the Tanta Headquarter, Egypt

ARTICLE INFO

ABSTRACT

An integrated spatio-temporal statistical, P-spline modeling, Mann—Kendall and Thiel—Sen trend detection, and factorial analyses techniques were performed on 438 chemical and 274 microbiological data collected from twenty drinking water supply wells over four years (2010—2013) in Tanta District (Egypt). The prime objective was to characterize the spatio-temporal quality trends using indicators of turbidity, pH, TDS, Cl^-, SO_4^{2-}, Na^+, Total Alkalinity, hardness (Total, Mg, and Ca), Fe^{2+}, Mn^{2+}, Al^{3+}, Cu^{2+}, Zn^{2+}, F^-, NH_4^+, NO_2^-, NO_3^-, PO_4^{3-}, SiO_2, bacterial, and algal contents. Factorial analysis was applied to identify the significant factors loading the quality variation. Out of the 20 wells, notable upward trends were significant (>95% level) for the total hardness (30%), total alkalinity (20%), TDS (15%), Fe^{2+} (15%), Mn^{2+} (15%), NO_2^- (10%), and 10% for the NH_4^+, PO_4^{3-}, and SiO_2. Attenuation rates (mg/l/year) were higher in NH_4^+ (av. 0.023) than Mn^{2+} (av. 0.013) and Fe^{+2} (av. 0.010), and remarkable average rates were 6.77 (TDS), 3.27 (total alkalinity), 2.12 (total hardness), 0.79 (SiO_2), 0.011 (PO_4^{3-}), and 0.002 (NO_2^-) in decreasing order. High precision of the trend estimate was confirmed for the Mn^{2+}, NH_4^+, and Fe^{2+} data. Five factors were found to explain 78.8% of the total variance of the quality variables and in particular, a significant load of hardness parameters, Total alkalinity, TDS, Mn^{2+}, NH_4^+, and Fe^{2+} in decreasing order were identified. The spatio-temporal variation in pollutants originated from organic matter degradation, either naturally from the aquifer peaty sediments or anthropogenic due to improper well head protection in the urban centers or from the agricultural drains in low relief areas. Our findings have important societal implications regarding the management of the limited and valuable water resources in arid and semi-arid lands. The adopted methodologies could be readily applied to similar alluvial aquifers elsewhere.

1. Introduction

Pollution control and remediation of groundwater quality is a worldwide essential prerequisite for mitigating water scarcity and sustainable development. Due to the rapid growth of human needs in many sectors, groundwater resources are always subjected to significant challenges accommodated by pollution and health hazard risks. Management of these limited resources is paramount in arid and semi-arid regions experiencing rapid urban development (Grimm et al., 2008). Alluvial regions are more accessible to pollution risks due to high population densities and intense agricultural and industrial activities. Understanding of the factors governing variation and detection of groundwater quality trends in these regions are essential for providing an early warning system for quality changes where protection and sound management of the resource can be efficiently set.

Spatio-temporal trend tests in groundwater quality studies are rare and difficult because the meteorological, hydrological, and anthropological sources of variation can act individually or combined and vary in intensity both in space and time. The joint modeling of both spatial and time elements in a single spatio-

temporal modeling framework becomes indispensable for more coherent interpretation of status and trends of site groundwater characteristics (Evers et al., 2015). Groundwater quality data influence the selection of a statistical trend analysis method and its ability to correctly recognize a monotonic trend. Most time series data inherit a possible existence of seasonality, skewness, serial correlation, non-normal data, "less-than" (censored) values, outliers and missing values (Hirsch and Slack, 1984; Helsel and Hirsch, 1992; Mozejko, 2012). Variability in time can be cyclical with the seasons, steadily (a trend), sudden jumps or some other established variations. The Mann−Kendall (M−K) and Spearman's rho tests (Mann, 1945; Kendall, 1938, 1975) and Thiel−Sen (T−S) test (Thiel, 1950; Sen, 1968) are widely used for detecting trends and their magnitude and for assessing future changes, respectively. The non-parametric Mann−Kendall test is commonly used in hydrological sciences as it is rather insensitive to outliers and has been proven as powerful as the Spearman's test (Yu et al., 2002). These tests have no distributional assumptions and missing data or outliers are permitted (Gibbons, 1994). Numerous recent examples include groundwater level variation (Panda et al., 2007; Moukana and Koike, 2008; Moukana et al., 2013), groundwater monitoring (Broers et al., 2009), hydroclimatology (Ehsanzadeh et al., 2012), and rainfall-groundwater dynamics (Machiwal and Jha, 2014).

Based on the above background, this study is aimed at assessing the spatio-temporal status and trends in groundwater quality in the rapidly urbanizing arid to semi-arid Tanta District (Egypt) adopting advanced and rarely applied spatial trend analysis techniques. This study represents a vital phase in the selection and design of an appropriate and reliable resource management system. Such assessment has been rarely addressed in Egypt despite of the significant socio-economic impacts from the impaired groundwater quality that can be amplified by the high anthropic pressure and the ineffective groundwater management. The adopted methodologies could be readily applied to similar worldwide alluvial aquifers.

2. Hydrogeology and health risks

2.1. Study area

Tanta district of the Gharbiya governorate (Egypt) is located in the heart of the delta midway between the Rashid and Damietta branches of the Nile River (Fig. 1). The district is a heavily cultivated agricultural site that consists of the capital city, Tanta, 50 villages, and 200 small hamlets. It covers an area of about 331 km^2 out of 1942 km^2 of the governorate area. It exhibits a Mediterranean climate with annual rainfall events of 50 mm mostly in winter. Groundwater supplies more than 85% of the drinking water, abstracted from the water-bearing formations of about 500 m thick Pleistocene and Holocene deposits. The Pleistocene aquifer (Mit Ghamr Formation) is composed mainly of Nilotic loose quartz sands and gravels interbedded with thin clay beds. It is generally capped by relatively thin characteristic silty cap beds sometimes intercalated with peat seams (Bilqas Formation of Holocene age). The Pleistocene aquifer is confined to semi-confined underlying a leaky phreatic aquifer formed of the Holocene deltaic deposits saturating the subsoil water.

In Tanta district, the upper 350 m of the 500 m Pleistocene aquifer is saturated with fresh water resting upon the deteriorated deeper saline water. According to Elewa (2010) and references therein, the leaky aquifer attains an average transmissivity of 15,000 m^2/day and hydraulic conductivity of 0.14 cm/s. The water level is shallow just 1−1.5 m below the ground surface. Vertical infiltration from irrigation canals and drainage water are the chief recharge to this aquifer. The inferred infiltration velocity of irrigation water averages 3 m/day in Tanta district (Elewa, 2010 and references therein). Replenishment from the underlying Pleistocene sands and gravel aquifer through the upward leakage is also possible. Groundwater flows northwards following the topography of flat-lying alluvial plain averaging 8.5 m above mean sea level ranging between 11 m at the south and 6 m in its northern part.

2.2. Health risks of polluted groundwater

Special emphasis is devoted to quality parameters exceeding their aesthetic guidelines including Mn^{2+}, Fe^{2+}, NH_4^+, and Bacteria. Mn^{2+} and Fe^{2+}, at higher concentrations in groundwater-sourced drinking water supplies, can cause adverse health effects or illness such as neurotoxin (McArthur et al., 2012a,b; Wasserman et al., 2006), hyperactive classroom behaviors (Bouchard et al., 2007), with children most vulnerable (Ljung and Vahter, 2007) and can hamper the intellectual development of the child (Buschmann et al., 2008). Because of the association between Mn^{2+} intake and neurological effects, the WHO current guideline concentration for Mn^{2+} in drinking water is 0.4 mg/L (World Health Organization, 2011). Iron and manganese tend to co-occur in groundwater (both are soluble only under oxygen-poor conditions) and hence, they yield similar indicators. In the oxygen-poor groundwater in which Fe^{2+} and Mn^{2+} are soluble; high NH_4^+ and low NO_3^- are common. Iron has been associated with genetic and metabolic diseases and repeated blood transfusions (Fraga and Oteiza, 2002). Higher concentrations of NH_4^+ are often considered to be sourced from human induced activities such as fertilizer application (manure), landfill leachate, effluent and wastewater discharges, amongst others (Böhlke et al., 2006; Kuroda et al., 2006). Although the usual level of NH_4^+ does not pose a direct risk to human health, its high content suggests the presence of more serious residential or agricultural contaminants, such as pathogens or pesticides. Bacteria in drinking water are major cause of waterborne diseases including diarrhea (the second biggest killer of children under five), hepatitis and typhoid (WHO, 1999). Microorganisms and bacteria in particular can dramatically accelerate the transport of heavy metals in the subsurface because bacterial surfaces are often negatively charged in natural environments (Bekhit et al., 2009 and references therein), and they can effectively bind dissolved cationic metals to their surfaces (e.g., Fowle and Fein, 1999), or indirectly alters the pH and Eh conditions of the microenvironment, thereby influencing Mn speciation and concentration (Nealson, 1983).

3. Data and methods

3.1. Data inventory and preprocessing

A groundwater quality-monitoring network is lacking in Tanta City and data from the public supply wells are alternatively used which are monitored on a regular basis as an integral part of the quality surveillance of national drinking water supply. Twenty three groundwater quality indicators of Turbidity, pH, TDS, Cl$^-$, SO$_4^{2-}$, Na$^+$, Total Alkalinity, hardness (Total, Mg, and Ca), Fe^{2+}, Mn^{2+}, Al^{3+}, Cu^{2+}, Zn^{2+}, F$^-$, NH$_4^+$, NO$_2^-$, NO$_3^-$, PO$_4^{3-}$, SiO$_2$, bacterial and algal contents had been collected for the four years from January 2010 to December 2013. Data at the twenty public water supply wells were obtained from the database of the Holding Company for Drinking Water and Wastewater (HCDWW) headquarter in the Tanta headquarter. Total counts of bacteria, total coliform, and algal samples were analyzed using the USEPA method (USEPA, 2000). This technique is fully approved as a procedure for

monitoring drinking water microbial quality in many countries (Rompré et al., 2002). All microbial data are reported as colony forming unit (CFU) per 100 mL. Sampling was carried out monthly but the frequency varied along the monitoring period in each well. In some months, samples are collected twice and in others there were missing observations. Therefore, the data were on a quarterly basis aggregated for trend analysis.

A major challenge of this study was sporadic water quality observations across space and time over a district, which leads to either lack of appropriate continuous temporal data or adequate spatial representation of the study area. The microbiological data values for these wells did not follow any regular "time-step" so as to provide robust estimate of trends and changes in contamination levels therein. A noticeable drop in available water quality observations over time was noticed for the third quarter in 2012. In view of the above mentioned limitations, water quality observations were aggregated on a quarterly basis for temporal analysis over the study area. To ascertain adequate spatial representation all available water quality observations for each quarter were used in the analysis. To accommodate for varying sample sizes over time, percentages of observations exceeding environmental guidelines were calculated, along with respective descriptive statistics, to support inferences drawn about trends on a quarterly basis in water quality.

A flow chart of the employed techniques is shown in Fig. 2. The site-specific descriptive statistics of the mean, median, minimum, maximum, standard deviation, and counts were performed using all available data for all the period. Comparisons to water quality standards are performed on a per-parameter basis, to determine the number and percentage of monitoring sites at which the means exceed the local permissible limits. Violating the threshold does not always indicate a threat to human health, because some guidelines are purely aesthetic, and in the case of health-related standards, water treatment methods can often be employed to remove or reduce the concentration of the contaminants. The data were not normally distributed, which is common in groundwater data (Harris et al., 1987; Helsel and Hirsch, 1992) and no general homogeneity of variance was
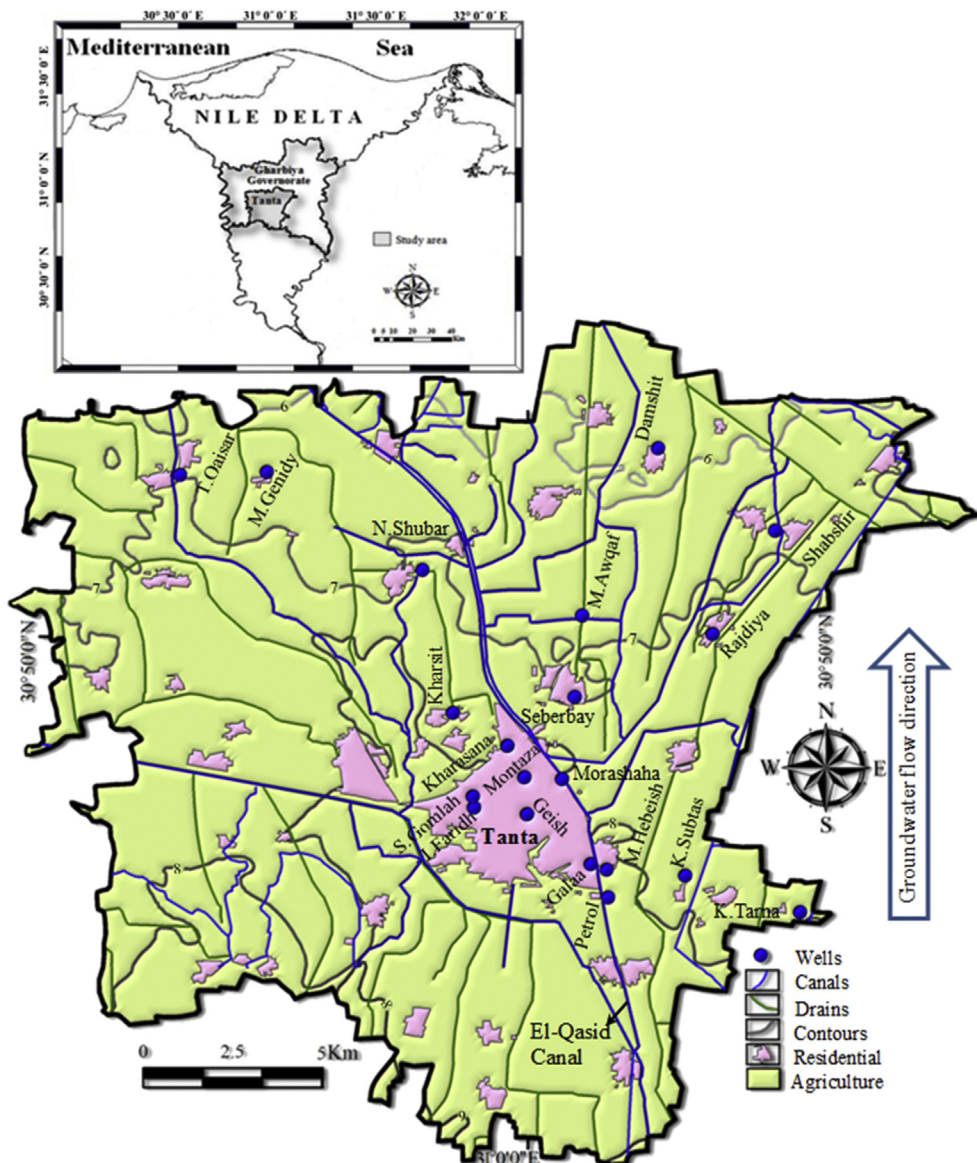

Fig. 1. Location map and physiographic features of Tanta District in the Middle Delta (Egypt).

observed. Therefore, non-parametric tests were used for further statistical analysis, as is also recommended for groundwater analysis by Helsel and Hirsch (2002).

3.2. Groundwater trend analysis

Trend assessments were performed for the 23 quality variables; trends for trace elements were not determined because they were analyzed at relatively few sites and the reported concentrations were mostly near or below the detection limit. Trend magnitude, the rate of change in each parameter, based on Sen's slope estimator for all trends are detectable with the Mann–Kendall test at the 95% confidence interval of increasing (positive), decreasing (negative) and non-significant trends. The Turbidity, pH, TDS, Cl^-, SO_4^{2-}, Na^+, Total Alkalinity, hardness (Total, Mg, and Ca), Fe^{2+}, Mn^{2+}, Al^{3+}, Cu^{2+}, Zn^{2+}, F^-, NH_4^+, NO_2^-, NO_3^-, PO_4^{3-}, SiO_2 were selected for the trend analysis, because their data were longest. Additionally, these are important indicators of specific processes: salinization (TDS, Cl^-, Na^+, and SO_4^{2-}); agricultural inputs (PO_4^{3-} and NO_3^-); acidification or water hardening (pH, total alkalinity, and total hardness), nutrients (NH_4^+), trace elements (Al^{3+}, Cu^{2+}, Zn^{2+}, and F^-), and redox potential (Fe^{2+} and Mn^{2+}). Trend analysis was performed with GWSDAT (Jones et al., 2015) and XLSTAT, Excel add-ons capable of performing a fully automatic trend analysis of a data set without special data requirements. Their effectiveness has been demonstrated extensively in the assessment of soil and groundwater conditions at Shells downstream assets (Jones et al., 2014). Tests determine if observed water quality changes over time are due to random variability or statistically significance. The nonparametric M–K and T–S trend statistical tests were used to detect and estimate trends and magnitude, respectively. The non-parametric M–K test for 95% level of significance monotonic trend was applied to the 14 groundwater quality indicators with long and near complete records. This test copes with autocorrelation and/or seasonality. The M–K test computes the slope of the line formed by plotting the variable of interest per time, but only considers the sign not the magnitude of the slope. M–K searches a trend in a time series without specifying whether the trend is linear or nonlinear (Khaliq et al., 2009). It firstly calculates the S statistic, a sum of the difference between the data points indicated by:

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} sgn(x_j - x_k) \tag{1}$$

where x_j and x_k are observed values at times j and k, $j > k$ ($j = 2, ..., n$; $k = 1, ..., n - 1$) and n is the total number of data. The value of S indicates the direction of trend. A negative (positive) value indicates falling (rising) trend. If $n \geq 8$, the S value is approximately normally distributed with mean and variance, $Var(S)$ as:

$$Var(S) = \frac{n(n-1)(2n+5) - \sum_{i=1}^{q} t_p(t_p - 1)(2t_p + 5)}{18} \tag{2}$$

where q is the number of tied groups (a set of sample data having the same value) and t is the value in the pth group. The statistical significance of S is checked using a test statistic (or Z score), which follows the standard normal distribution with zero mean and unit variance:

$$Z = \begin{cases} \dfrac{s-1}{\sqrt{Var(s)}} & ; S > 0 \\ 0 & ; S = 0 \\ \dfrac{s+1}{\sqrt{Var(s)}} & ; S < 0 \end{cases} \tag{3}$$

If the value of $|Z| > z_{\alpha/2}$, the null hypothesis (i.e., there is no distinct trend) is rejected at α level of significance in a two sided test. Kendall's τ represents a probability of the difference between the probabilities that the two variables are in the same order in the observed data versus the probability that the two variables are in different orders. Kendall's τ coefficient measures the strength of upward or downward trend. In this analysis, a 95% confidence level is used to define a statistically significant trend.

Linear regression is generally used to estimate the slope of possible linear trend in time series, but for noisy short time series data, it may grossly miscalculate the magnitude of trend. The Sen's slope estimator (mg/l/yr) was calculated to estimate trend magnitude (Sen, 1968, Hirsch et al., 1991). Sen's slope estimator used in this study is a robust non-parametric trend operator to determine the magnitude of the trend is highly immune to gross data error
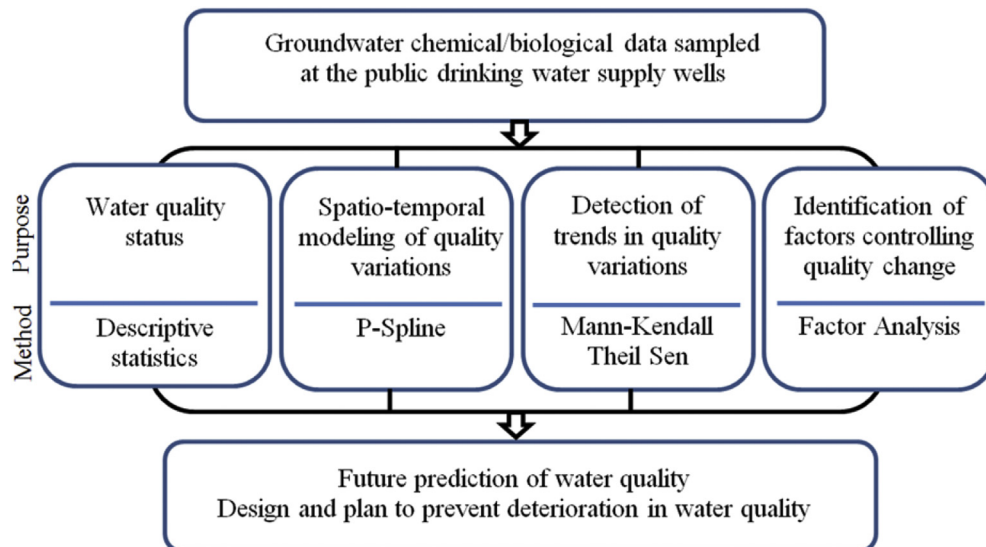

Fig. 2. Flow chart showing the methods adopted for investigating the status and trends of the groundwater quality indicators.

(Sen, 1968). The possible slopes between all possible data pairs are calculated:

$$Q_i = \frac{x_j - x_k}{j - k} \qquad (4)$$

where Q_i is the slope between data pair j and k ($i = 1 \dots N$; N is the total number of pairs, $N = n\,(n-1)/2$. The Sen's estimator for the median value of Q_i is:

$$Q = \begin{cases} Q_{\left[\frac{N+1}{2}\right]} & ; odd\ N \\ 0.5\left(Q_{\left[\frac{N}{2}\right]} + Q_{\left[\frac{N+2}{2}\right]}\right) & ; even\ N \end{cases} \qquad (5)$$

3.3. Spatio-temporal modeling of solute concentrations

The spatio-temporal solute concentration is smoothed using Penalized Splines (P-Splines) (Eilers and Marx, 1992, 1996), a nonparametric regression technique because this can simultaneously estimate both the spatial and time series trend in site solute concentrations. The fit of the spatio-temporal algorithm to the monitoring data were then evaluated using local polynomial regression techniques. More details are described in Evers et al. (2015). The following outlines some of the aspects of the technique. Let y_i be a natural log solute concentration at $x_i = (x_{i1}, x_{i2}, x_{i3})$ where x_{i1} and x_{i2} stand for the spatial coordinates of the well and x_{i3} represents the corresponding time for the i-th observation with $i = 1, \dots, n$. The first step is a modeling of the solute concentration as Equation (6)

Table 1
Summary statistics of wells with commonuality variables exceeding the permissible limits (PL).

	PL	1	0.3	0.4	0.5		PL	1	0.3	0.4	0.5
		Turb	Fe^{2+}	Mn^{2+}	NH$_4^+$			Turb	Fe^{2+}	Mn^{2+}	NH$_4^+$
I.Faridh	Av.	0.60	0.23	**0.91**	**1.24**	Damshit	Av.	**2.12**	**0.33**	**0.43**	0.32
	Min.	0.22	0.05	0.29	0.38		Min.	1.52	0.14	0.33	0.15
	Max.	1.65	0.45	2.3	3.28		Max.	2.74	0.49	0.53	0.55
	STD.	0.29	0.11	0.70	0.90		STD.	0.49	0.09	0.06	0.12
	Count.	24	23	23	24		Count.	14	13	14	13
Petrol	Av.	**2.69**	**0.65**	**0.66**	**0.91**	Seberbay	Av.	0.34	0.11	**0.42**	0.17
	Min.	1.08	0.18	0.1	0.25		Min.	0.1	0.02	0.32	0.03
	Max.	5.94	2.55	0.91	1.34		Max.	0.78	0.23	0.6	0.45
	STD.	1.47	0.70	0.25	0.33		STD.	0.13	0.07	0.07	0.13
	Count.	16	16	16	15		Count.	22	15	22	20
Galaa	Av.	0.28	0.15	**0.74**	**0.98**	Shabshir	Av.	**1.62**	**0.35**	**0.43**	0.28
	Min.	0.02	0.02	0.05	0.03		Min.	0.24	0.06	0.05	0.05
	Max.	1.75	0.53	1.83	2.17		Max.	4.21	0.55	0.7	0.54
	STD.	0.44	0.14	0.56	0.74		STD.	0.91	0.13	0.18	0.13
	Count.	20	12	20	19		Count.	27	26	26	24
Geish	Av.	0.21	0.05	**0.40**	0.28	N.Shubar	Av.	**1.48**	0.29	**0.40**	0.26
	Min.	0.02	0.02	0.06	0		Min.	0.15	0.14	0.21	0.08
	Max.	0.92	0.08	0.8	0.65		Max.	3	0.53	0.53	0.58
	STD.	0.26	0.03	0.19	0.18		STD.	0.79	0.10	0.07	0.14
	Count.	18	8	16	18		Count.	23	23	22	23
Kharasana	Av.	**1.40**	**0.47**	**0.50**	0.42	K.Tarna	Av.	**1.25**	0.23	**0.45**	0.26
	Min.	0.12	0.05	0.01	0.08		Min.	0.43	0.04	0.05	0.02
	Max.	5.8	1.89	1.21	1.12		Max.	2.74	0.65	0.63	0.42
	STD.	1.80	0.49	0.25	0.27		STD.	0.60	0.13	0.16	0.11
	Count.	23	18	22	23		Count.	21	20	18	20
S.Gomlah	Av.	**1.23**	**0.38**	**0.51**	**0.79**	M.Hebeish	Av.	**1.00**	0.16	**0.47**	0.34
	Min.	0.09	0.05	0.2	0.3		Min.	0.22	0.05	0.32	0.06
	Max.	2.62	0.76	0.92	1.61		Max.	5.7	0.29	0.65	1.25
	STD.	0.52	0.17	0.13	0.25		STD.	1.33	0.07	0.09	0.35
	Count.	20	20	19	20		Count.	15	13	14	11
T.Qaisar	Av.	**1.88**	**0.42**	**0.87**	**1.58**	M.Genidy	Av.	**1.91**	**0.46**	0.36	**0.52**
	Min.	0.54	0.12	0.25	0.88		Min.	1.17	0.29	0.25	0.38
	Max.	5.12	0.98	1.44	2.24		Max.	2.63	0.71	0.4	0.74
	STD.	1.07	0.19	0.27	0.38		STD.	0.56	0.17	0.06	0.14
	Count	16	16	16	15		Count.	5	5	5	5
Kharsit	Av.	**1.29**	**0.34**	**0.4**	**0.60**	All	Av.	**1.07**	**0.31**	**0.47**	**0.5**
	Min.	0.1	0.03	0.04	0.28		Min.	0.02	0.02	0.01	0
	Max.	3.17	0.59	0.54	1.71		Max.	5.94	2.55	2.3	3.28
	STD.	1.30	0.18	0.13	0.31		STD.	1.03	0.25	0.32	0.52
	Count.	20	14	19	21		Count.	388	337	366	367

NB: PL is the permissible limit for the qualityvariables. NTU is the unit for turbidity and mg/L for Fe^{2+}, Mn^{2+}, and NH$_4^+$.The symbols Av., Min., Max., STDV, and Count stand for the average, minimum, maximum, standard deviation, and the count of variables of the analyzed samples.Values exceeding the PL are highlighted in pink color.

Table 2
Number (NS) and percentage (%) of samples exceeding (Exc.) the permissible limits for the common quality indicators.

	TNS	Turbidity			Mg Hardness			Fe^{2+}			Mn^{2+}			NH$_4^+$			NO$_2^-$		
		NS	Exc.	%	NS	Exc.	%	NS	Exc.	%	NS	Exc.	%	NS	Exc.	%	NS	Exc.	%
I.Faridh	29	29	1	3.45	29	4	14	28	7	25	28	20	**71**	28	23	**82**			
Petrol	16	16	16	**100**	16	3	19	16	13	**81**	16	13	**81**	15	13	**87**			
Galaa	31	31	4	13	31	2	6	20	1	5	31	14	**45**	26	17	**65**			
Geish	19				19	2	11				19	6	**32**	18	1	6			
Kharasana	24	24	10	**42**	24		0	18	11	**61**	23	16	**70**	24	9	**38**			
Morashaha	31	31	1	3				30	2	7									
Montaza	39	39	3	8	39	2	5	24	1	4	32	4	13				37	1	3
S.Gomlah	20	20	14	**70**				20	16	**80**	19	18	**95**	20	18	**90**			
Rajdiya	18	18	10	**56**				18	5	28	16	7	**44**				16	1	6
T.Qaisar	18	18	13	**72**				18	13	**72**	18	17	**94**	17	16	**94**			
Kharsit	20	20	9	**45**				14	10	**71**	19	12	**63**	21	10	**48**	16	1	6
Damshit	14	14	14	**100**				13	11	**85**	14	9	**64**		2				
Seberbay	22				22	2	9				22	13	**59**						
Shabshir	31	28	23	**82**				27	20	**74**	26	16	**62**	25	1	4	17	1	6
N.Shubar	23	23	18	**78**	23	1	4	23	12	**52**	22	14	**64**	23	2	9			
K.Subtas	19	19	2	11				19	7	**37**									
K.Tarna	21	21	14	**67**				20	2	10	18	13	**72**						
M.Awqaf	22	22	9	**41**				21	4	19	22	3	14				12	1	8
M.Hebeish	14	14	3	20							14	12	**86**	11	2	18			
M.Genidy	5	5	5	**100**				5	4	**80**	5	1	20	5	2	**40**	3	1	**33**
Total	436	393	169	**43**	203	16	8	334	139	**42**	364	208	**57**	233	116	**50**	101	6	6

NB: TNS is the total number of samples, NS is the number of samples with measured and detected parameters. Percentages exceeding 32% of the total samples where quality indicators violated the permissible limits are shown in bold and highlighted in pink.

$$y_i = \sum_{j=1}^{m} b_j(x_i)\alpha_j + \varepsilon_i \qquad (6)$$

where the b_j, $j = 1, \ldots, m$ are m B-Spline basis functions. The measurement errors ε_i are assumed to be independent and identically normally distributed with zero mean and variance σ^2. Rewriting Equation (6) in a compact matrix notation is

$$\mathbf{y} = \mathbf{B(x)}\alpha + \boldsymbol{\varepsilon} \qquad (7)$$

The traditional ordinary least squares approach is to minimize

Table 3
Number (NS) and percentage (%) of samples exceeding (Exc.) the permissible limits (PL) for the bacterial and algal contents.

	NS	Bacteria cfu/ 100 ml								Algae (Count / ml)						
		Total Count			Total Coliform			Fecal Streptococci		Brown		Green				Blue-green
		PL=50	%	Exc.	PL=2	%	Exc.	Exc.	%		%	Colony	%	single-celled	Filamentous	
I.Faridh	6									1	**17**					
Petrol	5															
Galaa	29	11	**38**	1				3	**10**	5	**17**	2	**7**	1		1
Geish	11	5	**45**		1	**9**	1			1	**9**					
Kharasana	16	8	**50**							2	**13**	1	**6**		1	
Morashaha	18	4	**22**							4	**22**	1	**6**			
Montaza	26	12	**46**	1	1	**4**	1			7	**27**	2	**8**			
S.Gomlah	5				2	**40**	1									
Rajdiya	15	8	**53**							1	**7**					
T.Qaisar	18	8	**44**	3	2	**11**	2									
Kharsit	17	4	**24**	1						4	**24**					
Damshit	11	5	**45**													
Seberbay	14	7	**50**													
Shabshir	17	8	**47**	1	2	**12**	2									
N.Shubar	18	5	**28**		1	**6**	1			2	**11**			1		
K.Subtas	15	5	**33**							1	**7**					
K.Tarna	17	7	**41**		4	**24**	4			1	**6**					
M.Awqaf	16	8	**50**							2	**13**					
M.Hebeish	6	3	**50**							5	**83**					
M.Genidy	5	2	**40**													
Total	274	110	**40**	7	13	**5**	12	3	**1**	36	**13**	6	**2**	2	1	1

the following objective function:

$$O\,(\alpha) = ||\varepsilon||^2 = ||\mathbf{y} - \mathbf{B(x)}\alpha||^2 \tag{8}$$

To overcome a problem of overfitting of the model to the data, the objective function is modified with adding a term that penalizes the finite differences of the coefficients of adjacent B-splines as:

$$O\,(\alpha) = ||\mathbf{y} - \mathbf{B(x)}\alpha||^2 + \lambda||\mathbf{D}_d\alpha||^2 \tag{9}$$

where $\mathbf{D}_d$ is a matrix for the d-th differences of α, $\mathbf{D}_d = \Delta \mathbf{d}$, and λ is a nonnegative tuning parameter. By minimizing the new objective function for a given value of λ, we obtain the most likelihood parameters as:

$$\widehat{\alpha} = \left(\mathbf{B'B} + \lambda\mathbf{D}'_d\mathbf{D}_d\right)^{-1}\mathbf{B'y} \tag{10}$$

The optimal λ value can be defined a Bayesian modeling framework (Evers et al., 2015).

3.4. Factor analysis

Factor Analysis (FA) models a set of observable variables in terms of a few common unobservable factors (of maximum variance) and the loading scores (correlation coefficients) of the variables on these factors. In FA, the principal components (PCs), linear combinations of the original variables and the eigenvectors (Wunderlin et al., 2001), are computed from covariance or cross-product matrix describing the dispersion of the multiple interrelated variables.

Let a set of p observable random variables be $x_1, \ldots, x_p$ with means $\mu_1, \ldots, \mu_p$. The deviation from the mean is supposed to be a linear combination of k unknown constants l_{ij} and unobserved random variables F_j, where $i \in 1, \ldots, p$, $j \in 1, \ldots, k$, and $k < p$, as

$$x_i - \mu_i = l_{i1}F_1 + \ldots + l_{ik}F_k + \varepsilon_i \tag{11}$$

or in the matrix form:

$$\mathbf{x} - \mathbf{\mu} = \mathbf{LF} + \mathbf{\varepsilon} \tag{12}$$

The $\mathbf{F}$ and $\mathbf{L}$ are termed factor and loading matrices. The ε_i are independently distributed errors with zero mean, $E(\varepsilon) = 0$, finite variance, $Var(\varepsilon_i) = \psi_i$, and covariance

$$Cov(\mathbf{\varepsilon}) = Diag(\psi_1, \ldots, \psi_p) = \mathbf{\Psi} \tag{13}$$

To obtain $\mathbf{F}$ and $\mathbf{L}$ from n observations, four assumptions are used: $\mathbf{L}$ is constant across observations, $\mathbf{F}$ and ε are independent, $E(\mathbf{F}) = 0$, and Cov $(\mathbf{F})$ is unit matrix so that the factors are uncorrelated with each other.

First few factors explaining the highest variance are placed foremost, interpreted and linked to a specific hydrochemical process. FA has been used to evaluate groundwater quality (Suk and Lee, 1999; Masoud, 2014), and to identify anthropogenic impacts versus background (Helena et al., 2000; Pereira et al., 2003). XLSTAT, an Excel statistical add-on was used for the FA.

4. Results and discussion

4.1. Physico-chemical groundwater quality

Descriptive summary statistics of the groundwater quality data that exceed the national guidelines is shown in Table 1. Out

Table 4
Pearson's correlation matrix of groundwater quality variables.

	Turbidity.	pH	TDS	Cl$^-$	SO$_4^{2-}$	Na$^+$	Total alkalinity	Total hardness	Fe^{2+}	Mn^{2+}	Al^{3+}	Cu^{2+}	Zn^{2+}	F$^-$	NH$_4^+$	NO$_2^-$	NO$_3^-$	PO$_4^{3-}$	SiO$_2$
Turbidity	1																		
pH	-0.07	1																	
TDS	0.03	-0.16	1																
Cl$^-$	0.06	-0.21	**0.91**	1															
SO$_4^{2-}$	-0.02	-0.17	**0.74**	**0.66**	1														
Na$^+$	-0.11	-0.14	**0.87**	**0.78**	**0.65**	1													
T. alkalinity	0.08	-0.01	**0.83**	**0.65**	**0.50**	**0.65**	1												
T. hardness	0.16	-0.11	**0.92**	**0.84**	**0.69**	**0.69**	**0.84**	1											
Fe^{2+}	**0.61**	0.03	0.00	-0.03	-0.02	0.00	0.06	0.11	1										
Mn^{2+}	0.01	0.08	**0.47**	**0.36**	**0.35**	**0.33**	**0.51**	**0.57**	-0.01	1									
Al^{3+}	-0.26	**0.41**	-0.11	-0.28	**0.44**	-0.21	**0.31**	-0.04	0.05	**0.61**	1								
Cu^{2+}	-0.05	0.02	0.13	0.06	0.15	0.01	0.20	0.11	-0.08	-0.09	0.29	1							
Zn^{2+}	0.25	0.15	-0.05	-0.04	-0.09	0.02	0.02	-0.06	**0.62**	-0.09	**0.92**	-0.09	1						
F$^-$	-0.14	0.11	0.18	0.19	0.21	**-0.32**	0.18	0.11	**0.35**	0.23	**0.74**	0.02	**0.75**	1					
NH$_4^+$	0.10	0.16	**0.32**	0.17	0.27	0.11	**0.46**	**0.46**	0.17	**0.75**	-0.06	0.12	-0.02	-0.08	1				
NO$_2^-$	0.01	0.04	0.00	-0.07	0.00	0.00	0.00	-0.02	0.03	-0.02	-0.18	0.20	0.23	**0.79**	-0.07	1			
NO$_3^-$	-0.25	-0.13	**0.38**	0.27	**0.35**	**0.51**	0.26	0.23	-0.14	-0.10	-0.21	0.01	0.05	-0.04	-0.12	0.09	1		
PO$_4^{3-}$	0.08	0.06	-0.05	-0.12	-0.01	-0.07	0.05	0.00	0.18	0.16		0.01	-0.09	0.12	0.22	0.19	-0.09	1	
SiO$_2$	-0.07	0.13	0.10	0.06	0.19	0.18	0.08	0.08	-0.03	0.22		**0.50**	0.12	0.10	0.17	-0.03	0.00	0.06	1

NB: Bold indicates significant values exceeding a correlation coefficient of 0.3.

of the twenty wells, the mean concentrations at 55%, 40%, 70%, 35%, 35%, and 25% of the wells are higher than the guidelines for turbidity (1 NTU), Fe^{2+} (0.3 mg/L), Mn^{2+} (0.4 mg/L), and NH_4^+ (0.5 mg/L), Total counts of bacteria (50 cfu/100 mL), and Coliform bacteria (2 cfu/100 mL), respectively. One well exceeded the guideline for Total and Ca hardness (Kharasana) and Mg hardness (Geish). The highest mean concentrations of turbidity, Fe^{2+}, Mn^{2+}, and NH_4^+ were recorded in Petrol (2.69), Petrol (0.65 mg/L), I. Faridh (0.91 mg/L), and T. Qaisar (1.65 mg/L), respectively. In the whole area, 57%, 50%, 43%, 42%, 8%, and 6% of the samples exceeded the guidelines for Mn^{2+}, NH_4^+, turbidity, Fe^{2+}, Mg hardness (150 mg/l), and Nitrite (0.2 mg/l), respectively (Table 2). More than 32% of the samples exceeded the Mn^{+2} guideline in 16 wells. NH_4^+ guideline has been violated in more than 38% of the

samples in 8 wells. The permissible limit for turbidity has been violated 100% of the samples in Petrol, Damshit, and M. Genidy wells. More than 37% of the samples exceeded the guidelines of Fe^{2+} at ten wells. Wells with oxygen-poor groundwater in which Fe^{2+} and Mn^{2+} are high and soluble tend to have high NH_4^+ and low NO_3^- dominated in S. Gomlah, Petrol, Kharasana, T. Qaisar, and Kharsit (Tables 1 and 2). Excess of dissolved Mn^{2+} and Fe^{2+} in groundwater is strongly related to the amount of organic matter naturally occurred in the aquifer sediments (Rotiroti et al., 2013) or significant inflow of sewage water from broken or leaking sewer pipes that facilitates the reductive dissolution of Mn- and Fe-bearing solids in aquifers (Chae et al., 2008). In such aquifers, NH_4^+ is released from the degradation of organic nitrogen of peat (Rotiroti et al., 2013).
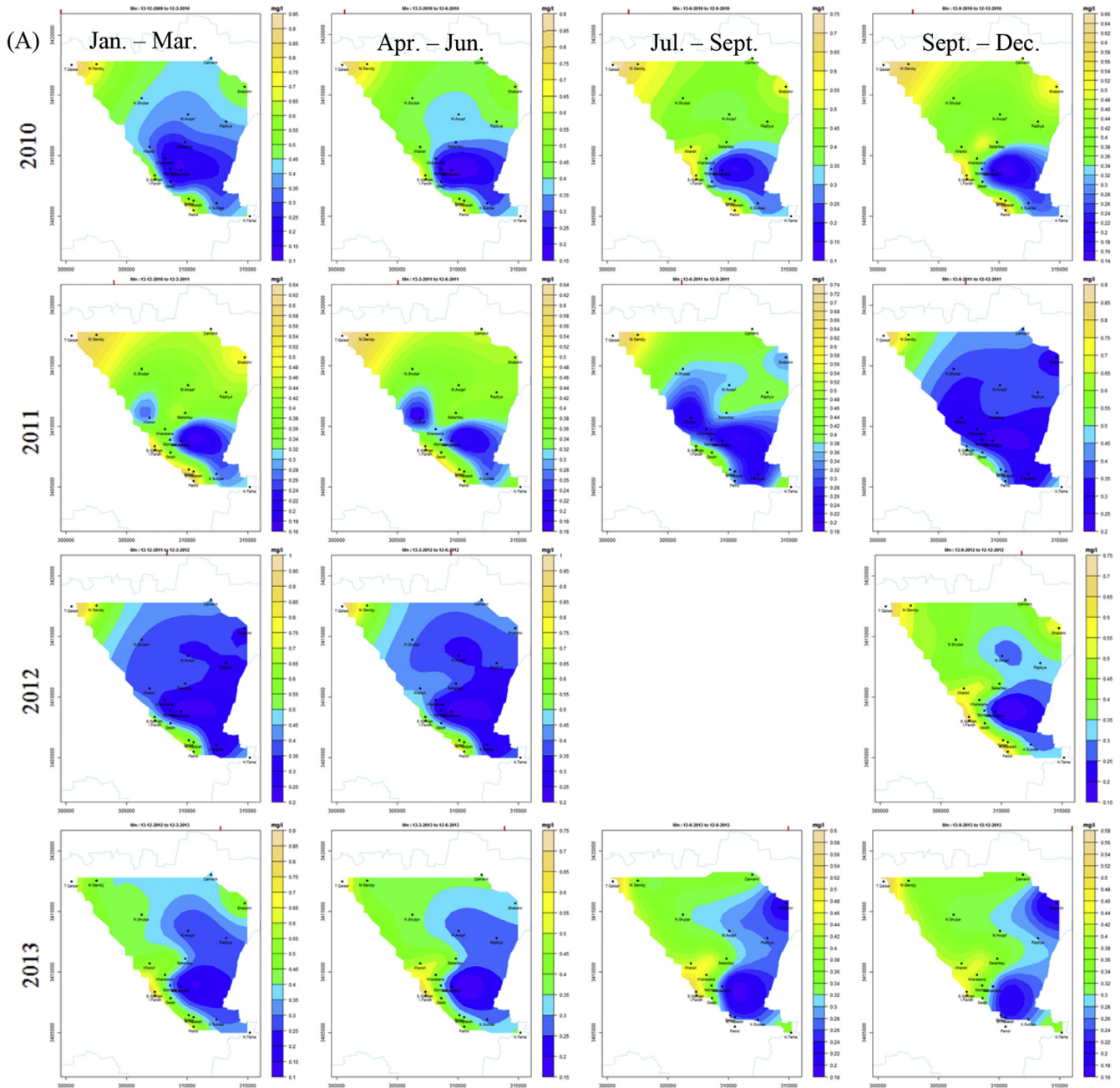

Fig. 3. Spatio-temporal variations of three ion concentrations by P-Spline, (A) Mn^{2+}, (B) NH_4^+, and (C).

4.2. Bacterial and algal contamination

The occurrence of total coliforms is interpreted as an indication of contaminated surface water infiltration and the potential occurrence of fecal pathogens (Payment and Locas, 2011). Of the 274 samples, 110 were contaminated against 164 absences of microbiological contamination (Table 3). In the whole area, the total counts (TC) of bacteria reached 40% of the samples. TC ranged from 22% to 50% of the samples in all wells while bacteria were not detected in the samples of I. Faridh, Petrol, and S. Gomla. TC exceeded the permissible limits (50 cfu/100 mL) in 7 samples of five wells. Total coliform violated the guidelines (2 cfu/100 mL) in 13 samples (5%) in 7 wells. Fecal coliform was not detected in the samples. Fecal Streptococci have been recorded only in 3 samples of Galaa indicating contamination by animal feces. K. Tarna recorded the most microbiologically contaminated well where the total coliform recorded in 4 samples out of 17 samples. Coliform usually indicates the human- or animal-originated contamination. High bacteria counts generally indicate contaminated groundwater. Sources of the contamination could include improperly functioning septic systems, poor well-head protection, and runoff from barnyards in permeable soils and short residence time.

Although there are no guidelines for the algal contents, the brown algae are the most common type, detected in 13 wells, and varied from 7% to 83% of the samples. Brown algae were of frequent occurrence in the samples of M. Hebeish (83%) followed by Montazah (27%), Kharsit (24%), and Morashaha (22%). Green algae of colony type existed in 4 wells of Montaza (8%), Galaa (7%), Kharasana (6%), and Morashaha (6%). Single-celled green
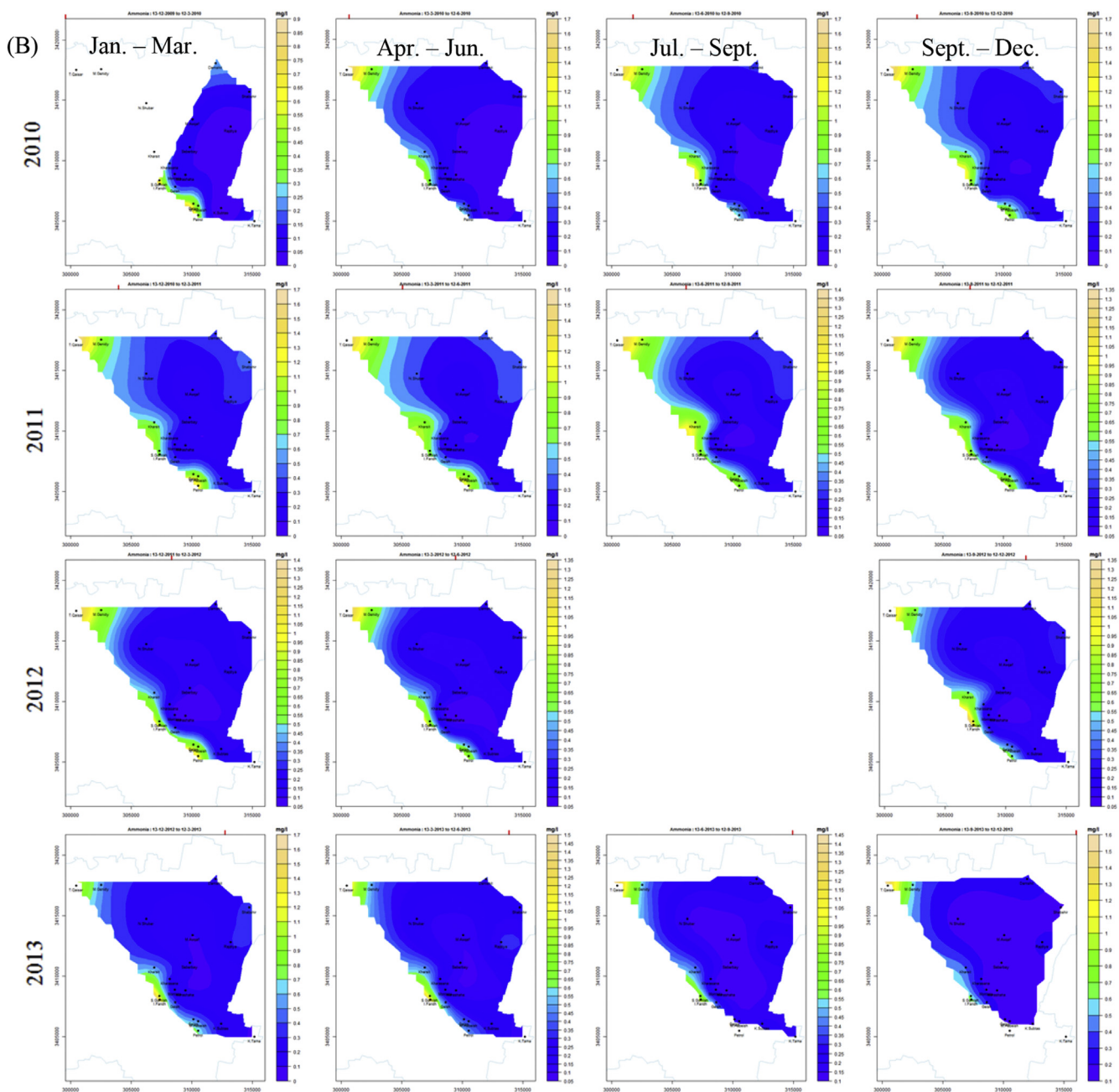

Fig. 3. (continued).

algae have been detected once in the samples of Galaa and N. Shubar. Filamentous green algae occurred in one sample of the Kharasana.

4.3. Correlation between quality variables

Pearson's correlation analysis was applied to the data (Table 4). Significant positive correlations among various indicators were evident. Turbidity is mainly related to the Fe^{2+} content ($r = 0.61$) agitated mostly during abstraction. TDS has significant correlation with total hardness, Cl^-, Na^+, total alkalinity, SO_4^{2-}, Mn^{2+}, NO_3^-, and NH_4^+ with $r = 0.74$ to 0.92 and low to moderate correlation $= 0.32$ to 0.47, with Mn, NO_3^-, and NH_4^+. This reflects the fact that TDS of groundwater is strongly related

to their contents. Mn^{+2} clarified strong correlation with NH_4^+ ($r = 0.75$) and moderate correlation ($r = 0.61$) with total alkalinity (as CO_3^{2-}), and did not correlate with Fe^{2+}. Therefore, distributions of the dissolved Fe^{2+} and Mn^{2+} constituents are not necessarily coincident as a consequence of redox zonation where reducing conditions progress with depth and starts by the reduction of Fe^{+2} followed by SO_4^{2-} and then Mn^{2+} (BGS, 2014). The aquifers consisted of organic-rich peaty sediment as the study area enhance Mn^{+2} mobilization through complexation with organic acids (humic or fulvic acids). The low nitrate concentrations in most wells are in agreement with suboxic conditions due to the presence of Mn^{2+} (e.g., Buschmann et al., 2007).

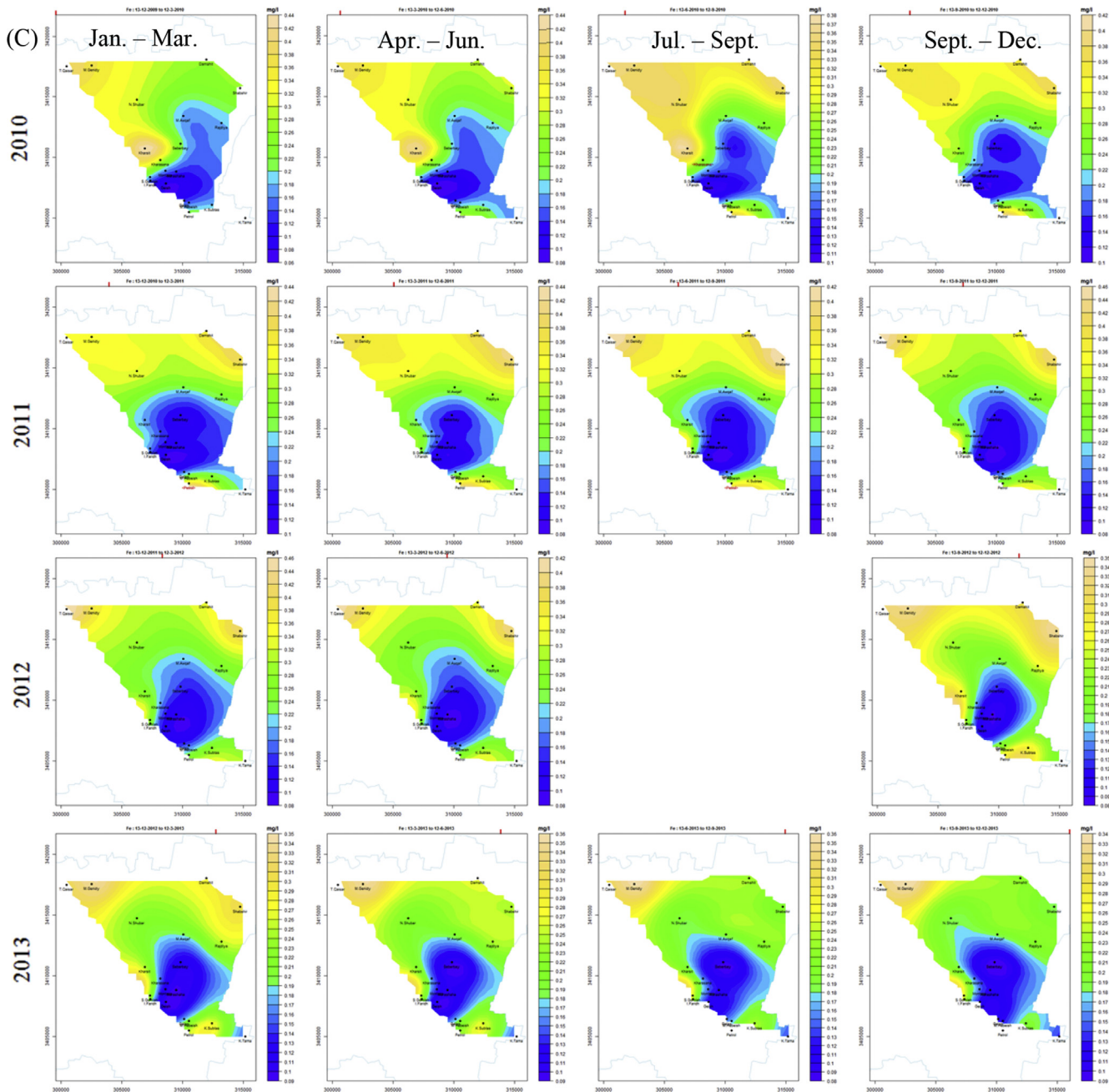The total alkalinity is attributed to natural processes such as
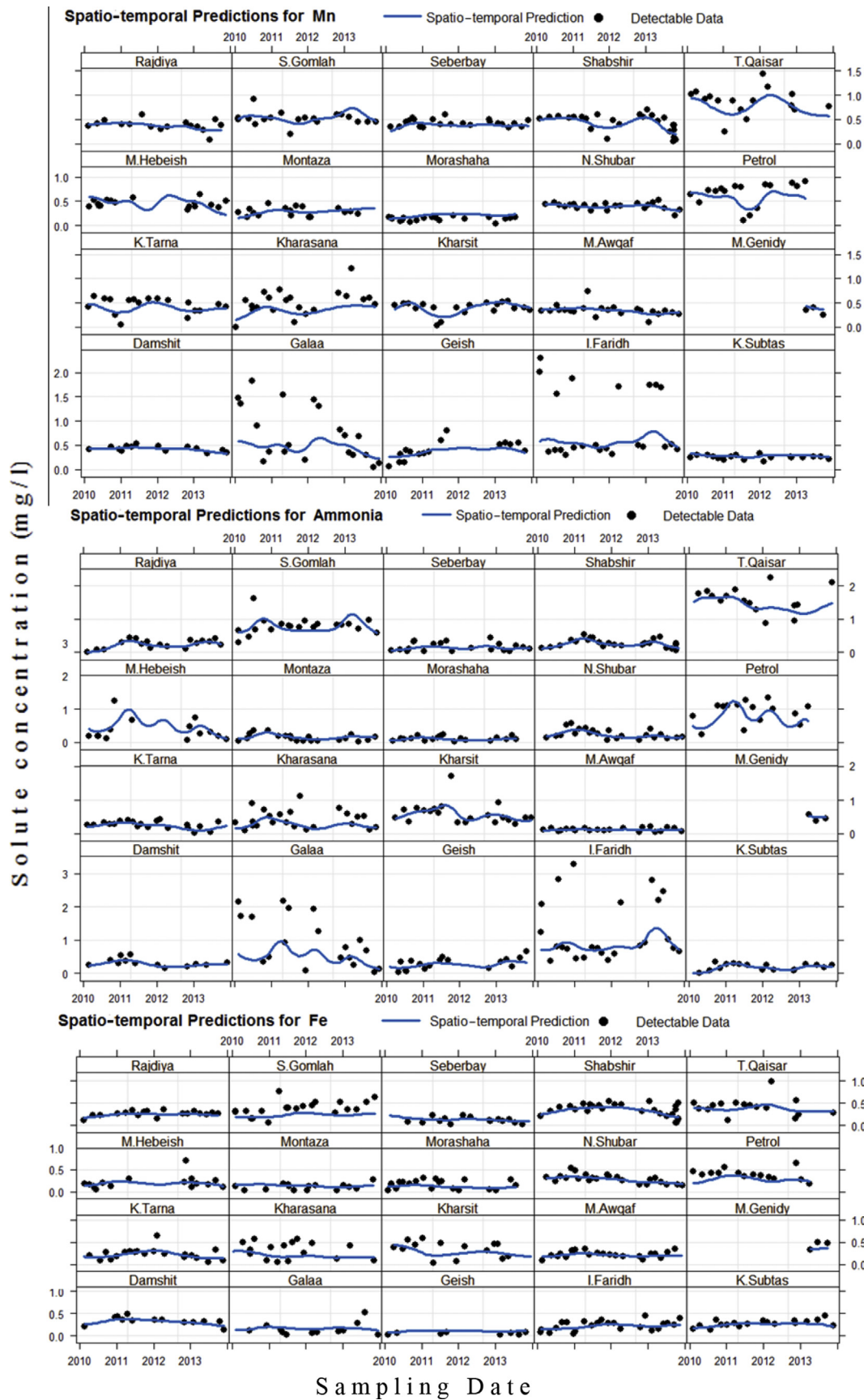

Fig. 3. (continued).

Fig. 4. Measured solute concentrations of Mn^{2+}, NH_4^+, and Fe^{2+} versus sampling date, with the fitted spatio-temporal model overlaid.

Table 5
Coefficient of determination of the linear relationship between predicted and measured concentrations of Mn^{2+}, NH_4^+, and Fe^{2+}.

Well	Solute			Well	Solute		
	Mn^{2+}	NH_4^+	Fe^{2+}		Mn^{2+}	NH_4^+	Fe^{2+}
I. Faridh	0.12	**0.40**	0.18	Damshit	**0.44**	**0.62**	**0.65**
Petrol	0.19	0.21	0.19	Seberbay	0.10	0.25	0.18
Galaa	0.12	0.22	0.08	Shabshir	**0.56**	**0.63**	**0.50**
Geish	**0.62**	**0.30**	0.19	N. Shubar	**0.37**	**0.52**	**0.57**
Kharasana	**0.45**	0.19	0.19	K. Subtas	0.12	**0.49**	0.22
Morashaha	0.19	0.22	0.18	K. Tarna	**0.30**	**0.43**	**0.38**
Montaza	0.18	**0.43**	0.16	M. Awqaf	0.21	0.12	**0.42**
S. Gomlah	0.19	0.19	0.22	M. Hebeish	0.13	**0.45**	0.22
Rajdiya	0.21	**0.73**	**0.42**	M. Genidy	**0.31**	**0.36**	0.24
T. Qaisar	**0.50**	**0.37**	**0.32**	Total	**0.44**	**0.64**	**0.43**
Kharsit	**0.47**	**0.30**	0.11	Pearson's coeff.	0.66	0.8	0.65

NB: Bold indicates significant values exceeding a correlation coefficient of 0.3.

dissolution of carbonate minerals in presence of soil CO_2. According to Sankar et al. (2014), the elevated dissolved carbonate content (total alkalinity as CO_3^{2-}) controls the solubility of Mn^{+2} through the precipitation of rhodochrosite ($MnCO_3$). Fe^{2+} correlated moderately with Zn^{2+} ($r = 0.62$) and weakly with F^- ($r = 0.35$) indicating that iron content plays a significant role in the dissolution and distribution of Zn^{2+} and F^- in the groundwater.

4.4. Spatio-temporal modeling of groundwater quality

The spatio-temporal models of the Mn^{2+}, NH_4^+, and Fe^{2+} by P-Splines (Fig. 3A–C) show no consistent temporal trend. However, their local highs were more widespread in the peripheral parts of

the area in 2010 and 2011 as compared to 2012 and 2013 data. Local highs of Mn^{+2} and NH_4^+ marked the T. Qaisar, S. Gomlah, and Petrol wells and the surrounding areas. Fe showed local highs at T. Qaisar, Kharsit, Shabshir, Petrol, and K. Subtas. At T. Qaisar and Shabshir, the ground surface elevation is lowest among the wells and has agricultural drains where pollutants were accumulated during its northward flow across the district. The whole area centered on Mit. Hebeish, Morashah, and Seberbay wells highlight the lowest Mn^{2+}, NH_4^+, and Fe^{2+} contents. The spatial distribution of these solutes may be related to the density and hence the recharge from the surface canals. In the central area with low solute concentrates, the largest canal, namely El-Qasid Canal, flows from south to north across the district and upon which Mit. Hebeish well is built (Fig. 1). All the quality indicators of this well are below the guidelines and show the best predicted quality among the studied wells in the future. Groundwater from this well has been extensively abstracted for bottled drinking water in Tanta and in Egypt.

4.5. Validation of the spatio-temporal quality modeling

Measured solute concentrations of Mn^{2+}, NH_4^+, and Fe^{2+} versus sampling date, overlaid with the fitted spatio-temporal model are shown in Fig. 4. The spatio-temporal models do not fit well with the sampling points at the wells of I. Faridh, Galaa, Kharasana, Morashaha, Montazah, and S. Gomla in two or more of Mn^{2+}, NH_4^+, and Fe^{2+}. The results of regression and correlation analyses were applied to quantify the variation of the variables in the studied wells by plotting the predicted against the measured values (Table 5). The whole area displayed a coefficient of determination (R^2) of 0.64, 0.44, and 0.43, and Pearson's coefficients (r) of 0.8, 0.66,
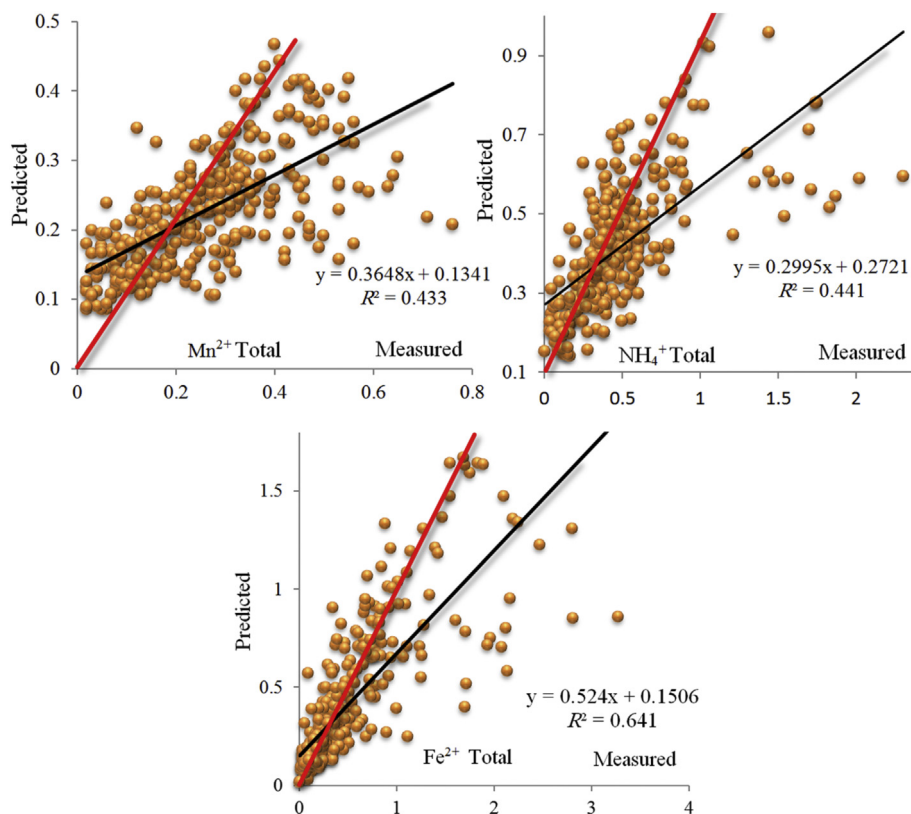

Fig. 5. Regression of predicted versus measured solute concentrations of Mn^{2+}, NH_4^+, and Fe^{2+}. Red line is the 45° trend line. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

and 0.65, for the Mn^{2+}, NH_4^+, and Fe^{2+}, respectively in decreasing order (Fig. 5). The maximum R^2 values are recorded at Geish (0.62), Rajdiya (0.73), and Damshit (0.65) in Mn^{2+}, NH_4^+, and Fe^{2+}, respectively. R^2 exceeding 0.3 for Mn^{2+}, NH_4^+, and Fe^{2+} was recorded at Shabshir, Damshit, T. Qaisar, N. Shubar, and K. Tarna. Good to moderate matching of the prediction with the measured data in these wells is mostly related to the continuous and long temporal record of the solute concentrations with smallest standard deviation. Another feature is the location in the peripheral part of the studied area in less densely-populated villages surrounded by arable lands, which minimizes the spatio-temporal variation in the data.

On the other hand, the wells of I. Faridh, Galaa, Kharasana, Morashaha, Montazah, and S. Gomla showed the weakest $R^2 < 0.2$ in two or more of the Mn^{2+}, NH_4^+, and Fe^{2+}. These wells have long records with the largest standard deviations because they are located in the densely populated zones. The urban areas are associated mostly with excess of sewage effluents and industrial wastes that vary extremely in time and space.

4.6. Physico-chemical trends

Ten quality variables of the studied wells revealed statistically significant increasing M–K trends (Table 6). Out of the 20 wells, upward trends were notable for the total hardness (30%), total alkalinity (20%), TDS (20%), Fe^{2+} (15%), Mn^{2+} (15%), NO_2^- (10%), and 5% for the NH_4^+, PO_4^{3-}, and SiO_2. The causes of upward trends in total hardness, TDS and alkalinity are probably related to an increase in non-point source pollution from changing land uses, along with use of fertilizers in arable land, and sewage effluents. Wells of Petrol, Geish, and Kharasana showed upward trends in Total alkalinity, TDS, Total hardness, and Mn^{+2}. In addition, the Geish well recorded upward trends of NH_4^+ and NO_2^-, suggesting the worst groundwater conditions in the near future (Fig. 6). The NH_4^+ will continue to increase based on the upward trend at the Geish well where the four-year NH_4^+ data exceeded the regulation value. Upward trends were recorded in Fe^{2+} (I. Farid, S. Gomlah, and K. Subtas), NO_2^- (M. Awqaf), PO_4^{3-} (Galaa), and SiO_2 (I. Farid).

Notable downward trends were found in turbidity (35%), TDS (20%), Total hardness (15%), Total alkalinity (15%), Mn^{2+} (15%), Fe^{2+} (10%), and NH_4^+ (5%). The best groundwater qualities are expected from the M. Hebeish, K. Tarna, and Damshit wells with downward trends in Turbidity, total hardness, total alkalinity, TDS and no significant trends were detected in other variables. Galaa well showed downward trends in TDS, Total hardness, Mn^{2+}, and NH_4^+ against an upward trend in PO_4^{3-}. Trend analysis of K. Tarna expected good quality conditions with no upward trends in chemical variables despite the most microbiologically contaminated well.

Attenuation rates of change (mg/l/year) showed average of 6.77 (TDS), 3.27 (total alkalinity), 2.12 (total hardness), 0.79 (SiO_2), 0.011 (PO_4^{3-}), and 0.002 (NO_2^-) in decreasing order. Rates of change (mg/l/year) were higher in NH_4^+ (*av.* 0.023) than Mn^{2+} (*av.* 0.013) and Fe^{+2} (*av.* 0.01). Geish well showed the highest rate of increase followed by Petrol and Kharasana wells for TDS (14.16, 8.47, and 5.5), total alkalinity (6.66, 5, and 2.5), and total hardness (4.85, 4, and 6.66), respectively.

The results of the quality statistics, exceedance of the permissible limits, correlations between quality variables, and the sustained trends of the quality indicators clarify emerging problems of Mn^{2+}, NH_4^+, and Fe^{+2}. Elevated NH_4^+ is mostly associated with high Mn^{2+} and lower NO_3^- at large depths. Trend analysis indicated significant upward/downward trends for TDS, total alkalinity, total hardness, SiO_2, Mn^{2+}, Fe^{2+}, NO_2^-, and NH_4^+. Rates of change (mg/L/yr) vary between 14.16 and 0.007 for TDS and Fe^{2+}, respectively. Either the individual or the combined effects of variation in relief, density and proximity to the reaches of agricultural drainage/surface canals, intensive use of fertilizers in arable land, and sewage

Table 6
Statistically significant increment/decrement trends at 95% with Kendall τ and S, and Sen's slope.

	Turbidity			Total alkalinity			TDS			Total hardness		
	τ	S	Slope	τ	S	Slope	τ	S	Slope	τ	S	Slope
Petrol	−0.60	−72	−0.2	**0.87**	**104**	**5**	**0.76**	**91**	**8.47**	**0.41**	**48**	**4**
Galaa							*−0.38*	*−64*	*−9.3*	*−0.34*	*−65*	*−6.15*
Geish				**0.57**	**96**	**6.66**	**0.5**	**85**	**14.16**	**0.45**	**76**	**4.85**
Kharasana				**0.34**	**85**	**2.5**	**0.36**	**90**	**5.5**	**0.48**	**119**	**6.66**
Morashaha	−0.36	−92	−0.018							**0.41**	**101**	**1.2**
Damshit	−0.66	−59	−0.8	*−0.421*	*−37*	*−0.8*	*−0.51*	*−46*	*−1.5*			
Seberbay										**0.4**	**91**	**4.62**
Shabshir	−0.34	−118	−0.059							**0.32**	**91**	**0.54**
N. Shubar	−0.47	−118	−0.078									
K. Subtas				**0.43**	**61**	**0.66**	**0.4**	**59**	**0.66**			
K. Tarna				*−0.493*	*−100*	*−1.5*	*−0.75*	*−158*	*−7.57*	*−0.61*	*−126*	*−7.46*
M. Awqaf	−0.38	−87	−0.036							*−0.45*	*−126*	*−10*
M. Hebeish	−0.48	−50	−0.046	*−0.410*	*−43*	*−8*	*−0.56*	*−58*	*−17.25*	*−0.45*	*−47*	*−10*

	Fe^{2+}			Mn^{2+}			NH_4^+			NO_2^-			PO_4^{3-}			SiO_2		
	τ	S	Slope	τ	S	Slope	τ	S	Slope	tau	S	Slope	τ	S	Slope	τ	S	Slope
I. Faridh	**0.26**	**65**	**0.007**													**0.51**	**34**	**0.79**
Petrol	*−0.44*	*−52*	*−0.014*	**0.44**	**52**	**0.015**												
Galaa				*−0.46*	*−88*	*−0.054*	*−0.43*	*−74*	*−0.011*				**0.38**	**39**	**0.011**			
Geish				**0.57**	**67**	**0.03**	**0.52**	**80**	**0.023**	**0.54**	**73**	**0.002**						
Kharasana				**0.22**	**49**	**0.01**												
S. Gomlah	**0.432**	**81**	**0.014**															
Shabshir				*−0.37*	*−118*	*−0.01*												
N. Shubar	*−0.48*	*−120*	*−0.01*															
K. Subtas	**0.49**	**68**	**0.01**															
M. Awqaf				*−0.34*	*−78*	*−0.004*				**0.52**	**33**	**0.0005**						

NB: Trends of increment are shown in bold and decrement is shown in italic. The symbols τ, S, and slope indicate the strength of the slope, the direction of the trend slope, and attenuation rate, respectively.
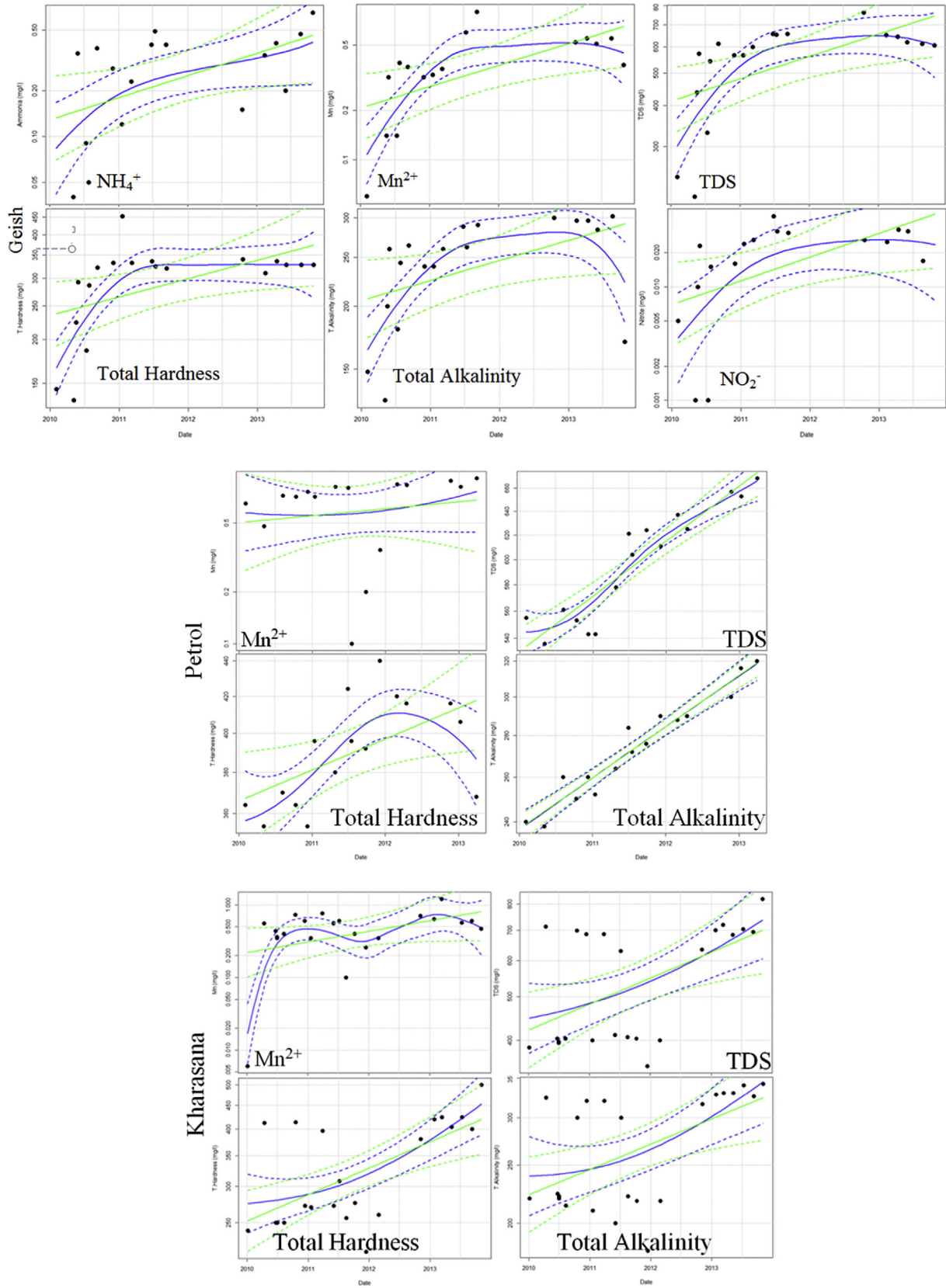
Fig. 6. Linear (solid green) and smooth (solid blue) trends at a confidence level of 95% (dashed green and blue) for wells attained four or more significant increasing quality parameters (Geish, Petrol, and Kharasana). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

Table 7
Factor analysis of the studied physic-chemical groundwater quality indicators.

	F1	F2	F3	F4	F5	Comm.	Var.
T. hardness	**0.97**	0.24	0.01	0.01	0.03	1.00	0.00
Ca hardness	**0.96**	0.24	0.05	−0.01	0.00	1.00	0.00
Mg. hardness	**0.95**	0.20	−0.08	0.07	0.11	1.00	0.00
T. alkalinity	**0.93**	0.02	−0.17	−0.24	0.08	1.00	0.00
TDS	**0.92**	0.37	−0.10	0.05	0.06	1.00	0.00
Mn^{2+}	**0.86**	−0.35	−0.20	−0.27	−0.16	1.00	0.00
NH_4^+	**0.59**	−0.55	−0.26	−0.07	−0.27	0.80	0.20
Fe^{2+}	**0.54**	−0.42	0.23	0.48	0.08	0.76	0.24
Turbidity	0.45	−0.33	**0.77**	−0.25	0.18	1.00	0.00
PO_4^{3-}	0.32	−0.04	**0.66**	0.19	−0.50	0.82	0.18
Zn^{2+}	0.27	−0.13	−0.24	**0.55**	0.08	0.45	0.55
F^-	0.16	0.10	−0.23	0.02	**0.36**	0.22	0.78
NO_2^-	0.08	**0.48**	−0.05	0.39	−0.17	0.43	0.57
NO_3^-	0.00	**0.68**	−0.16	−0.21	−0.33	0.64	0.36
Al^{3+}	−0.03	−0.50	**−0.67**	0.03	−0.20	0.74	0.26
Eigenvalue	6.7	2.2	1.8	1.2	0.7		
Variability (%)	41.7	13.9	11.2	7.5	4.4		
Cumulative %	41.7	55.7	66.8	74.4	78.8		

NB: Values in bold correspond to each variable to the factor for which the squared cosine is the largest.

effluents, are the major factors controlling spatio-temporal water quality changes.

4.7. Spatio-temporal variability factors

Due to the redundancy of information related to correlations existing among the variables, the first five factors with eigenvalues exceeding 78.8% of the total variance (Table 7) are suitable to characterize the data structure. Variables attained the largest correlation coefficients of the selected factors are interpreted. In addition, the communalities of the variables, the proportion of their variance explained by the extracted common factors, are larger than 0.6 and therefore represents a unique contribution to the discrimination of the processes controlling the variability in the

area. Therefore, FA is assumed to adequately represent the overall variance of the data set.

F1 contributes to 41.7% of the total variability and is significantly loaded by the hardness parameters, Total alkalinity, TDS, Mn^{2+}, NH_4^+, and Fe^{2+} in decreasing order of influence. F2 counts for 13.9% of the variability loaded by the NO_3^- and NO_2^- contents. F3 counts for 11.2% of the variability, positively loaded by the turbidity and the PO_4^{3-} content and negatively loaded by Al content. The association of the PO_4^{3-} and Al^{3+} contents with turbidity in F3 indicates a mutual causative relationship mostly originated from the excessive application of the phosphate fertilizers. F4 is dominated by Zn^{2+} control of 7.5% on variability. Finally, F5 is loaded by the F content and contributed to 4.4% of the total variability. Spatial distribution of variables defined by F1 versus F2 of the factor loadings (Fig. 7) displayed observable association of turbidity with the redox (Mn^{2+} and Fe^{2+}), nutrient (NH_4^+), trace elements (Zn^{2+}), and fertilizer input (PO_4^{3-}). TDS is closely associated with hardness and alkalinity parameters as well as with the anthropogenic NO_3^- and NO_2^- contents. In conclusion, groundwater hardness, alkalinity, salinity, and redox potentials are specified as the dominant indicators controlling the variations in groundwater quality.

5. Conclusion

This paper demonstrated usefulness of a combination of the parametric statistics, non-parametric Mann–Kendall and Theil–Sen trend analyses, P-Spline modeling, and the multivariate factorial analysis to clarify the status and the trends of spatio-temporal variations in the groundwater quality by a case study using the four-years (2010–2013) data sampled at the public supply wells in Tanta, Egypt. The chemical composition of the groundwater revealed the control of the water quality by several processes such as salinization, water hardness and alkalinity, redox process (Mn^{2+} and Fe^{2+}), and bacterial/algal inputs.

The descriptive statistics of the groundwater resources in Tanta revealed contrasting and emerging problems of Mn^{2+}, NH_4^+, and
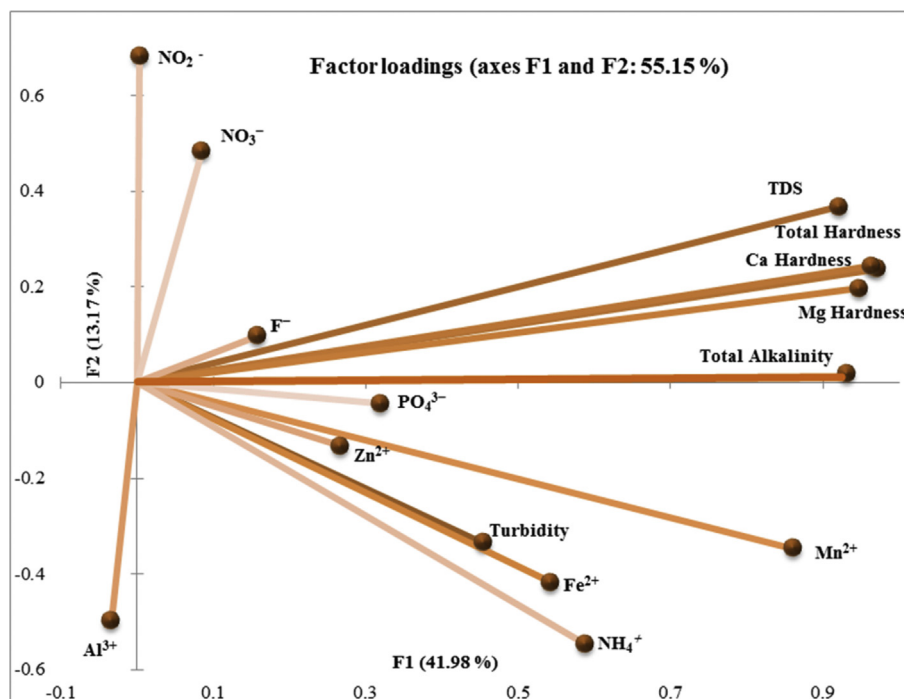

Fig. 7. Spatial distribution of variables in the space defined by F1 versus F2 of the factor loadings.

Fe^{+2} with the national drinking water limit exceeding at 57%, 50%, and 42% of the 20 sites sampled, respectively. Out of the 20 wells, trend analysis clarified notable upward trends for the total hardness (30%), total alkalinity (20%), TDS (15%), Fe^{2+} (15%), Mn^{2+} (15%), NO_2^- (10%), and 5% for NH_4^+, PO_4^{3-}, and SiO_2. Results showed average rate (mg/l/year) of increase over the four-year period higher in NH_4^+ (av. 0.023) than Mn^{2+} (av. 0.013) and Fe^{+2} (av. 0.004). Rates showed average of 6.77 (TDS), 3.27 (total alkalinity), 2.12 (total hardness), 0.79 (SiO_2), 0.011 (PO_4^{3-}), and 0.002 (NO_2^-) in decreasing order. Factor analysis showed that the five factors explained more than 78.8% of the total variance. A significant load of hardness parameters, Total alkalinity, TDS, Mn^{2+}, NH_4^+, and Fe^{2+} in decreasing order of influence, contributed to 41.7% of the total variability, while NO_3^- and NO_2^- contents to 13.9%. F3 counted for 11.2% of the variability, positively loaded by the turbidity and the PO_4^{3-} content and negatively loaded by Al^{3+} content. F4 is dominated by Zn^{2+} (7.5%) and F^- content contributed to 4.4% on the fifth factor.

Excess of the dissolved Mn^{+2} and Fe^{+2} is likely related to the degradation of the organic matter in the aquifer sediments (i.e., peat) or from significant sewage inflow that facilitates the reductive dissolution of Mn- and Fe-bearing solids in aquifers. NH_4^+ is mostly supported by the bacterial activities that alter the pH and Eh conditions under the micro-environment. The spatio-temporal differences in the pollutants originated from organic matter degradation, either naturally from of the aquifer peaty sediments or anthropogenic leaked due to improper well head protection in densely populated urban centers or from the proximity to agricultural drains in low relief areas. Identification of hydrologically similar zones underscores the necessity for further sub-division of management areas and implementation of zone-specific groundwater protection and/or restoration strategies. Although the most polluted water violated the national guidelines that entail various health hazards, is expected to be still underway to such public supply wells, the use of these wells should be avoided. Unless human management intervenes to reduce delivery of point/non-point source of pollution in Tanta, contaminants will likely continue to accumulate in these groundwater aquifers. Indeed, increased vigilance and monitoring to identify the presence of bacteria and viruses in groundwater and their viability in active recharge areas will likely come at a cost to future water quality.

Being a major center for industrialization and rapid urban sprawl, localized events of human-induced contamination from municipal and industrial waste disposal will warrant site-specific groundwater vulnerability studies that deserve analysis of the synergistic effects on human health. These studies can contribute to establish cause–effect relationships and design appropriate remedial strategies. Continuous monitoring of the studied quality indicators, in particular Mn^{2+}, NH_4, Fe^{2+}, TDS, NO_3^-, and Escherichia coli, should be carried out at a regular (periodic) interval (ideally on a quarterly basis), on an on-going basis, and at a consistent network of monitoring sites, all of which have suitable well-head protection. Integration of environmental factors such as socio-economic activity, land use, hydrogeology would be necessary to regulate groundwater quality in Tanta and in the Middle Delta areas having similar groundwater quality impairment problems.

The advocated methods are reliable and cost-effective and could potentially be used for detecting spatio-temporal trends of the groundwater contaminations in similar alluvial aquifers. Our findings have important societal implications regarding the management of the limited and valuable water resources in arid and semi-arid lands. The successful application and refinement of the methodologies developed for this study should invite similar applications elsewhere. This study serves as a replicable model, especially for similar settings in the less studied parts of the world.

Acknowledgment

References

Bekhit, H.M., El-Kordy, M.A., Hassan, A.E., 2009. Contaminant transport in groundwater in the presence of colloids and bacteria: model development and verification. J. Contam. Hydrol. 108 (3–4), 152–167.

Böhlke, J., Smith, R.L., Miller, D.N., 2006. Ammonium transport and reaction in contaminated groundwater: application of isotope tracers and isotope fractionation studies. Water Resour. Res. 42 (5), W05411.

Bouchard, M., Laforest, F., Vandelac, L., Bellinger, D., Mergler, D., 2007. Hair manganese and hyperactive behaviors: pilot study of school-age children exposed through tap water. Environ. Health Perspect. 115 (1), 122–127.

British Geological Survay-BGS, 2014. Water Quality Fact Sheet: Manganese. Available at: http://www.bgs.ac.uk/downloads/start.cfm?id=1275 (last checked 20 June 2014).

Broers, H.-P., Visser, A., Chilton, J.P., Stuart, M.E., 2009. Assessing and aggregating trends in groundwater quality. In: Quevauviller, P., Fouillac, A.-M., Grath, J., Ward, R. (Eds.), Groundwater Monitoring. John Wiley & Sons, Ltd, Chichester, UK. http://dx.doi.org/10.1002/9780470749685.ch12.

Buschmann, J., Berg, M., Stengel, C., Sampson, M.L., 2007. Arsenic and manganese contamination of drinking water resources in Cambodia: coincidence of risk areas with low relief topography. Environ. Sci. Technol. 41 (7), 2146–2152.

Buschmann, J., Berg, M., Stengel, C., Winkel, L., Sampson, M.L., Trang, P.T.K., Viet, P.H., 2008. Contamination of drinking water resources in the Mekong Delta floodplains: arsenic and other trace metals pose serious health risks to population. Environ. Int. 34 (6), 756–764.

Chae, G.T., Yun, S.T., Choi, B.Y., Yu, S.Y., Jo, H.Y., Mayer, B., Kim, Y.J., Lee, J.Y., 2008. Hydrochemistry of urban groundwater, Seoul, Korea: the impact of subway tunnels on groundwater quality. J. Contam. Hydrol. 101 (1–4), 42–52.

Ehsanzadeh, E., van der Kamp, G., Spence, C., 2012. The impact of climatic variability and change in the hydroclimatology of Lake Winnipeg watershed. Hydrol. Process. 26 (18), 2802–2813.

Eilers, P.H.C., Marx, B.D., 1992. In: Fahrmeir, L., et al. (Eds.), Generalized Linear Models with P-splines in Advances in GLIM and Statistical Modelling. Springer, New York.

Eilers, P.H.C., Marx, B.D., 1996. Flexible smoothing with b-splines and penalties. Stat. Sci. 11 (2), 89–121.

Elewa, H.H., 2010. Potentialities of water resources pollution of the Nile River Delta, Egypt. Open Hydrol. J. 4, 1–13.

Evers, L., Molinari, D.A., Bowman, A.W., Jones, W.R., Spence, M.J., 2015. Efficient and automatic methods for flexible regression on spatio-temporal data, with applications to groundwater monitoring. Environmetrics. http://dx.doi.org/10.1002/env.2347 (in press).

Fowle, D.A., Fein, J.B., 1999. Competitive adsorption of metal cations onto two gram positive bacteria: testing the chemical equilibrium model. Geochem. Cosmochim. Acta 63 (19–20), 3059–3067.

Fraga, C.G., Oteiza, P.I., 2002. Iron toxicity and antioxidant nutrients. Toxicology 180 (1), 23–32.

Gibbons, R., 1994. Statistical Methods for Groundwater Monitoring. John Wiley & Sons, New York.

Grimm, N.B., Faeth, S.H., Golubiewski,, N.E., 2008. Global change and the ecology of cities. Science 319, 756–760.

Harris, J., Loftis, J.C., Montgomery, R.H., 1987. Statistical methods for characterizing ground-water quality. Groundwater 25 (2), 185–193.

Helena, B., Pardo, R., Vega, M., Barrado, E., Fernandez, J.M., Fernandez, L., 2000. Temporal evolution of groundwater composition in an alluvial aquifer (Pisuerga River, Spain) by principal component analysis. Water Res. 34 (3), 807–816.

Helsel, D.R., Hirsch, R.M., 1992. Statistical Methods in Water Resources. Studies in Environmental Science, vol. 49. Elsevier Science Publishers BV, Amsterdam (chapter 13).

Helsel, D.R., Hirsch, R.M., 2002. Statistical Methods in Water Resources. Chapter A3, Book 4, Hydrologic Analysis and Interpretation, Techniques of Water-Resources Investigations of the United States Geological Survey (USGS). USGS, Reston, VA, p. 510.

Hirsch, R.M., Slack, J.R., 1984. A nonparametric trend test for seasonal data with serial dependence. Water Resour. Res. 20 (6), 727–732.

Hirsch, R.M., Alexander, R.B., Smith, R.A., 1991. Selection of methods for the detection and estimation of trends in water quality. Water Resources Research

27, 803–813.

Jones, W.R., Spence, M.J., Bowman, A.W., Evers, L., Molinari, D.A., 2014. A software tool for the spatio-temporal analysis and reporting of groundwater monitoring data. Environ. Model. Softw. 55, 242–249.

Jones, W.R., Spence, M.J., Bonte, M., 2015. Analyzing groundwater quality data and contamination plumes with GWSDAT. Groundwater 53 (4), 513–514.

Kendall, M.G., 1938. A new measure of rank correlation. Biometrika 30, 81–93.

Kendall, M.G., 1975. Rank Correlation Methods. Charles Griffin, London.

Khaliq, M.N., Ouarda, T.B.M.J., Gachon, P., Sushma, L., St-Hilaire, A., 2009. Identification of hydrological trends in the presence of serial and cross correlations: a review of selected methods and their application to annual flow regimes of Canadian rivers. J. Hydrol. 368 (1–4), 117–130.

Kuroda, K., Fukushi, T., Takizawa, S., Aichi, M., Hayashi, T., Tokunaga, T., 2006. Groundwater contamination and its sources in the Tokyo Metropolitan Area. Yousui Haisui 48, 37–45 (in Japanese with English abstract).

Ljung, K., Vahter, M., 2007. Time to re-evaluate the guideline value for manganese in drinking water? Environ. Health Perspect. 115 (11), 1533–1538.

Machiwal, D., Jha, M.K., 2014. Characterizing rainfall–groundwater dynamics in a hard-rock aquifer system using time series, geographic information system and geostatistical modelling. Hydrol. Process. 28 (5), 2824–2843.

Mann, H.B., 1945. Non-parametric tests against trend. Econometrica 13, 245–259.

Masoud, A.A., 2014. Spatio-temporal evaluation of the groundwater quality in Kafr Al-Zayat District, Egypt. Hydrol. Process. 27 (20), 2987–3002.

McArthur, J.M., Sikdar, P.K., Nath, B., Grassineau, N., Marshall, J., Banerjee, D.M., 2012a. Sedimentological control on Mn, and other trace elements, in groundwater of the Bengal Delta. Environ. Sci. Technol. 46, 669–676.

McArthur, J.M., Sikdar, P.K., Hoque, M.A., Ghosal, U., 2012b. Waste-water impacts on groundwater: Cl/Br ratios and implications for arsenic pollution of groundwater in the Bengal Basin and Red River Basin, Vietnam. Sci. Total Environ. 437, 390–402.

Moukana, J.A., Koike, K., 2008. Geostatistical model for correlating declining groundwater levels with changes in land cover detected from analyses of satellite images. Comput. Geosci. 34 (11), 1527–1540.

Moukana, J.A., Asaue, H., Koike, K., 2013. Co-kriging for modeling shallow groundwater level changes in consideration of land use/land cover pattern. Environ. Earth Sci. 70 (4), 1495–1506.

Mozejko, J., 2012. In: Voudouris (Ed.), Detecting and Estimating Trends of Water Quality Parameters, Water Quality Monitoring and Assessment. http://dx.doi.org/10.5772/33052. ISBN: 978-953-51-0486-5, InTech, Available from: http://www.intechopen.com/books/water-quality-monitoring-and-assessment/detecting-and-estimating-trends-of-water-quality-parameters.

Nealson, K.H., 1983. Microbial cycle of manganese. In: Krumbein, W.E. (Ed.), Microbial 422 Geochemistry. Blackwell, Oxford, pp. 191–221.

Panda, D.K., Mishra, A., Jena, S.K., James, B.K., Kumnar, A., 2007. Influence of drought and anthropogenic effects on groundwater level in Orissa, India. J. Hydrol. 343 (3–4), 140–153.

Payment, P., Locas, A., 2011. Pathogens in water: value and limits of correlation with microbial indicators. Groundwater 49 (1), 4–11.

Pereira, H.G., Renca, S., Sataiva, J., 2003. A case study on geochemical anomaly identification through principal component analysis supplementary projection. Appl. Geochem. 18, 37–44.

Rompré, A., Servais, P., Baudart, J., de-Roubin, M.R., Laurent, P., 2002. Detection and enumeration of coliforms in drinking water: current methods and emerging approaches. J. Microb. Methods 49 (1), 31–54.

Rotiroti, M., Bonomi, T., Fumagalli, L., 2013. An integrated approach to assess origin and mobilization of As, Fe and Mn in groundwater: the case study of Cremona (northern Italy). Geophys. Res. Abstr. 15. EGU2013–11097.

Sankar, M.S., Vega, M.A., Defoe, P.P., Kibria, M.G., Ford, S., Telfeyan, K., Neal, A., Mohajerinc, T.J., Hettiarachchi, G.M., Barua, S., Hobsona, C., Johannesson, K., Datta, S., 2014. Elevated arsenic and manganese in groundwaters of Murshidabad, West Bengal, India. Sci. Total Environ. 488–489 (1), 570–579.

Sen, P.K., 1968. Estimation of regression co-efficients based on Kendall's tau. J. Am. Stat. Assoc. 63, 1379–1389.

Suk, H., Lee, K.K., 1999. Characterization of hydrochemical system through multivariate analysis: clustering into groundwater zones. Groundwater 37 (3), 358–366.

Thiel, H., 1950. A rank-invariant method of linear and polynomial regression analysis part 3. Proc. Koninalijke Nederl. eAkademie van Weinenschatpen A53, 1397–1412.

US Environmental Protection Agency, 2000. Membrane Filter Method for the Simultaneous Detection of Total Coliforms and Escherichia coli in Drinking Water. Publication EPA 600-R-00–13.

Wasserman, G.A., Liu, X., Parvez, F., Ahsan, H., Levy, D., Factor-Litvak, P., 2006. Water manganese exposure and children's intellectual function in Araihazar, Bangladesh. Environ. Health Perspect. 114 (1), 124–129.

WHO, 1999. The World Health Report, Making a Difference. World Health Organization, Gineva.

World Health Organization, 2011. Library Cataloguing-in-publication Data. Guidelines for Drinking-water Quality, fourth ed., pp. 1–518

Wunderlin, D.A., Diaz, M.P., Ame, M.V., Pesce, S.F., Hued, A.C., Bistoni, M.A., 2001. Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia river basin (Cordoba, Argentina). Water Res. 35 (12), 2881–2894.

Yu, P.S., Yang, T.C., Chou, C.C., 2002. Effects of climate change on evapotranspiration from paddy fields in southern Taiwan. Clim. Change 54 (1–2), 165–179.