



A review of GIS-integrated statistical techniques for groundwater quality evaluation and protection

Deepesh Machiwal¹ · Vincent Cloutier² · Cüneyt Güler³ · Nerantzis Kazakis⁴

Received: 2 April 2018 / Accepted: 20 September 2018 / Published online: 26 September 2018
© Springer-Verlag GmbH Germany, part of Springer Nature 2018

Abstract

Water quality evaluation is critically important for the protection and sustainable management of groundwater resources, which are variably vulnerable to ever-increasing human-induced physical and chemical pressures (e.g., overexploitation and pollution of aquifers) and to climate change/variability. Preceding studies have applied a variety of tools and techniques, ranging from conventional to modern, for characterization of the groundwater quality worldwide. Recently, geographic information system (GIS) technology has been successfully integrated with the advanced statistical/geostatistical methods, providing improved interpretation capabilities for the assessment of the water quality over different spatial scales. This review intends to examine the current standing of the GIS-integrated statistical/geostatistical methods applied in hydrogeochemical studies. In this paper, we focus on applications of the time series modeling, multivariate statistical/geostatistical analyses, and artificial intelligence techniques used for groundwater quality evaluation and aquifer vulnerability assessment. In addition, we provide an overview of salient groundwater quality indices developed over the years and employed for the assessment of groundwater quality across the globe. Then, limitations and research gaps of the past studies are outlined and perspectives of the future research needs are discussed. It is revealed that comprehensive applications of the GIS-integrated advanced statistical methods are generally rare in groundwater quality evaluations. One of the major challenges in future research will be implementing procedures of statistical methods in GIS software to enhance analysis capabilities for both spatial and temporal data (multiple sites/stations and time frames) in a simultaneous manner.

Keywords Artificial intelligence methods · Geostatistical modeling · GIS · Hydrogeochemistry · Multivariate analysis · Time series modeling · Water quality index

✉ Deepesh Machiwal
Deepesh.Machiwal@icar.gov.in; dmachiwal@rediffmail.com

Vincent Cloutier
Vincent.Cloutier@uqat.ca

Cüneyt Güler
cuneytguler@gmail.com

Nerantzis Kazakis
kazakis@geo.auth.gr

¹ ICAR-Central Arid Zone Research Institute, Regional Research Station, Bhuj, Gujarat 370105, India

² Groundwater Research Group, Institut de Recherche en Mines et en Environnement, Université du Québec en Abitibi-Témiscamingue, Amos, Canada

³ Mersin Üniversitesi, Çiftlikköy Kampüsü, Jeoloji Mühendisliği Bölümü, 33343 Mersin, Turkey

⁴ Laboratory of Engineering Geology and Hydrogeology, School of Geology, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

Introduction

Groundwater is the major source of freshwater supply worldwide, which is currently used to meet nearly half of the drinking water needs, including requirement of about two billion people all over the world (WWAP 2009; Gleeson et al. 2010). In addition, groundwater provides around 43% of the water consumed in irrigation (Siebert et al. 2010). Accordingly, the major groundwater systems of the world do not remain in dynamic equilibrium rather do show significantly declining groundwater level trends (WWAP 2012). It has been estimated that about 700–800 km³ of groundwater has been depleted from the aquifers in the USA during the twentieth century (Konikow and Kendy 2005). Likewise, a World Bank Report (World Bank 2010) states that India is the largest consumer of groundwater in the world, with an estimated annual groundwater use of 230 km³. The fast-depleting groundwater resources, as depicted by

the declining groundwater levels, caused deterioration in groundwater quality in many parts of the world. Groundwater quality may be further degraded due to pressure created over hydrologic and hydrogeologic systems in view of the impacts of climate change and variability (e.g., Gurdak et al. 2012; Bondu et al. 2016).

The degradation of groundwater quality has also led to the reduction of the exploitable quantities. There are two major sources of the groundwater quality degradation, i.e., natural (geogenic) processes and anthropogenic activities. For example, in agricultural areas, excessive use of fertilizers has resulted in nitrate contaminations in groundwater well above the water quality guidelines (e.g., Machiwal et al. 2011; Paradis et al. 2016). In coastal areas, the over-exploitation of aquifers via numerous wells and boreholes has established a negative water balance triggering seawater intrusion and salinization of groundwater (Ferguson and Gleeson 2012). Anthropogenic activities that pollute the natural environment with potential toxic elements (such as hexavalent chromium, arsenic, and antimony) include paint manufacturing, tannery industry, mining activities, phosphate fertilizer manufacturing and the combustion of coal and fly ash deposits (Molina et al. 2009; Jacobs and Testa 2004). Natural processes also influence groundwater quality at both local and regional scales depending on the geological, hydrochemical and hydrogeological regimes. The release of arsenic (As) in groundwater is mainly controlled by the oxidation of organic carbon coupled with the reductive dissolution of As-bearing iron oxides (Postma et al. 2012), while As concentrations exceeding the WHO provisional guideline value of 10 µg/L (WHO 2017) were measured in groundwater from fractured bedrock aquifer of the Canadian Shield, primarily derived from the weathering of As-bearing sulfides in the oxic/suboxic zone of the aquifer (Bondu et al. 2017). Additionally, geothermal fluids can also enhance As concentrations in shallow aquifers through deep fractures (Pique et al. 2010; Iskandar et al. 2012). Iskandar and Koike (2011) identified the deep-seated hydrothermal system as the major source of As contamination along the fault zone in North Sulawesi, Sulawesi Island, Indonesia, by applying geostatistical and numerical simulation models. The geogenic origin of hexavalent chromium is attributed to ophiolitic rocks and specifically their serpentized derivatives (Nriagu and Nieboer 1988), while the concentrations of hexavalent chromium are influenced by the prevailing hydrogeological conditions (Kazakis et al. 2015). It is very much needed to pay adequate attention for evaluation of the quality of this vital but invisible resource from local to regional scales based on scientific knowledge, to manage it in a sustainable manner. Therefore, throughout the world, increasing demands for safe drinking water, agricultural and industrial use of it, as well as maintaining healthy ecosystems are leading stakeholders and scientists to develop

appropriate strategies and methods varying from simple to complex nature for rational groundwater resource management and protection (e.g., Council of Canadian Academies 2009).

There exist several conventional tools and techniques ranging from graphical to statistical that have been used by various researchers for interpreting groundwater quality (Freeze and Cherry 1979; Karanth 1987; Sara and Gibbons 1991; Güler et al. 2002; Machiwal and Jha 2010). In recent times, researchers felt a need for application of the modern techniques such as time series modeling (e.g., trend identification), multivariate statistics, and geostatistical modeling, among others, to better interpret and precisely characterize the groundwater quality for the efficient management and protection of groundwater resources (e.g., Güler et al. 2002; Jha et al. 2007; Cloutier et al. 2008; Steube et al. 2009; Machiwal and Jha 2010, 2015). The modern techniques also contribute to distinguish between the anthropogenic and natural processes and/or factors influencing the groundwater quality. Salient popular methods used for groundwater quality evaluation and protection are classified into distinct groups and subgroups as shown in Fig. 1.

With the advent of geographic information system (GIS) technology, especially after 1990s, visualization, interpretation and presentation of groundwater quality evaluations over large spatial scales has been drastically improved. The GIS is capable of capturing, storing, analyzing, manipulating, retrieving and displaying a large volume of spatial data for swift organization, quantification and interpretation for decision-making in areas including engineering and environmental sciences (e.g., Stafford 1991; Goodchild et al. 1993; Burrough and McDonnell 1998; Lo and Yeung 2003). It has been proved to be a powerful tool for analyzing and mapping the hydrologic/hydrogeologic data over spatial and temporal scales in order to provide useful information about spatio-temporal variability that ultimately helps in decision-making (Burrough and McDonnell 1998; Gurnell and Montgomery 2000; Chang 2002; Chen et al. 2004; Güler and Thyne 2004a; Machiwal and Jha 2014). The GIS applications are advantageous in groundwater quality evaluation studies particularly for mapping spatial variations of water quality, subsurface flow and pollution modeling, and groundwater-quality monitoring network design, etc. (Jha et al. 2007). In addition, GIS-based water quality mapping is imperative for pollution-hazard modeling, assessment and protection planning, and detection of environmental changes (Goodchild et al. 1993; Skidmore et al. 1997; Chen et al. 2004; Jha et al. 2007).

Recently, advanced statistical tools are successfully integrated with GIS by researchers for illustrating spatial distribution of the chemical composition of groundwater over the large areas, sometimes up to regional scale. In addition, many studies have utilized the GIS-integrated statistical

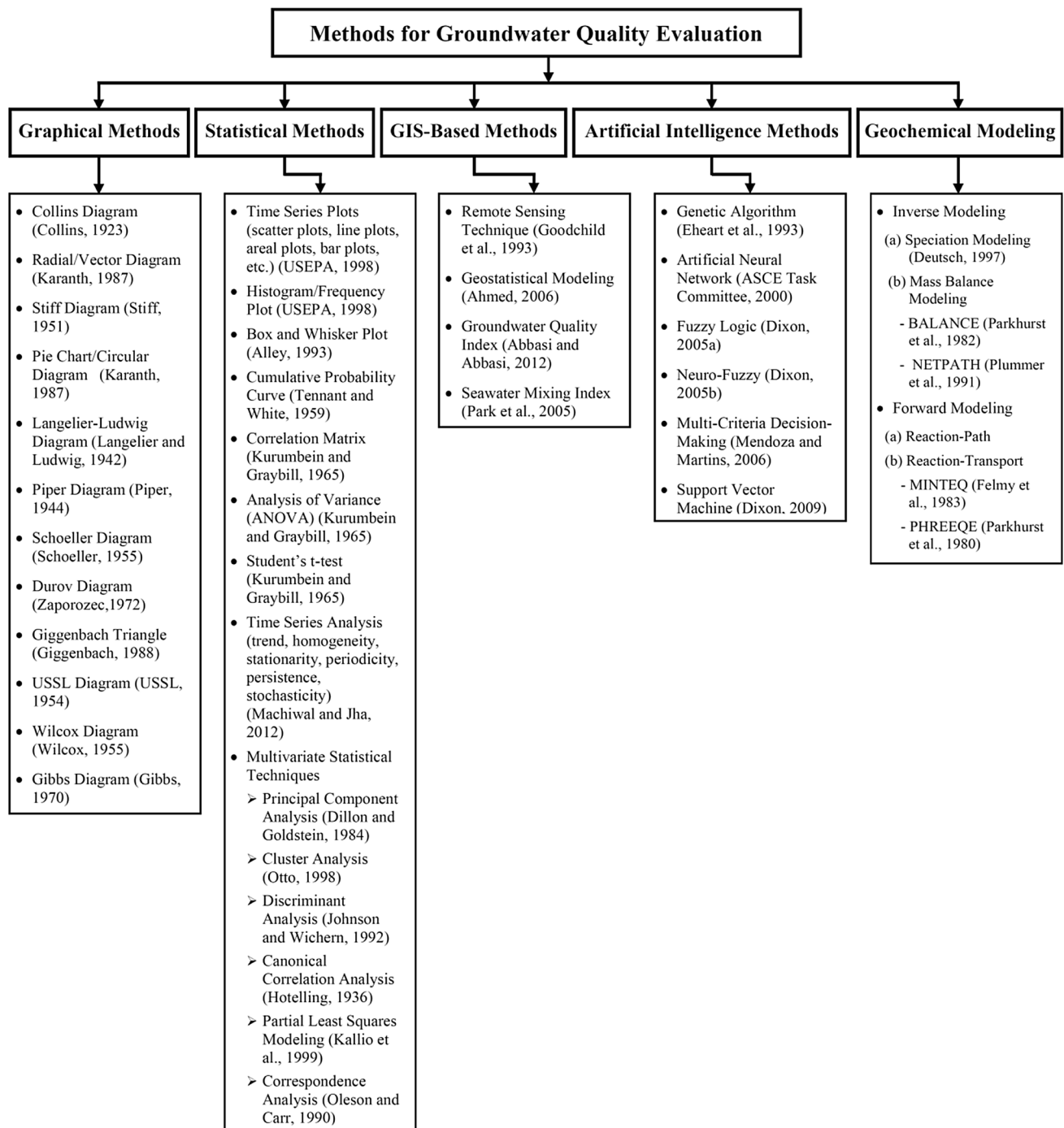


Fig. 1 Classification of salient methods for the groundwater quality evaluation and protection

techniques and approaches to determine the hydrochemical regime and establish strategies to manage groundwater resources under the complexity of natural processes in conjunction with the anthropogenic practices that influence groundwater quality. The purpose of this paper is to explore the literature in order to evaluate the current standing of the GIS-integrated statistical analyses used for groundwater

quality evaluation, and to focus on the future research directions. To the authors' knowledge, this kind of review does not currently exist in the literature. This paper deals with the past studies having application of time series modeling and multivariate statistical/geostatistical techniques for groundwater evaluation and artificial intelligence techniques for groundwater vulnerability assessment. Then,

this paper outlines distinct GIS-based water quality indices developed and applied for groundwater quality assessment worldwide. Finally, it highlights the limitations and research gaps incurred in the past studies, and then it emphasizes on the future research needs to be considered for a better evaluation of groundwater quality under the framework of GIS-integrated statistical techniques.

Time series modeling of groundwater quality variables

Components, steps and assumptions of time series analysis

A “time series” may be defined as “a sequence of values collected over time on a particular variable” (Haan 1977). Similar to the time-variable data series, there exist “spatial data series” in hydrogeology. In spatial series, data are location-dependent instead of time-dependent as in the time series. Most time series analysis techniques are equally applicable to spatial data series (Shahin et al. 1993), and hence, spatial data series is sometimes referred to as time series. In general, a hydrologic or hydrogeologic time series is composed of deterministic and stochastic components (Haan 1977; Shahin et al. 1993). The deterministic component presents a systematic pattern in the time series and can be classified as a trend, a shift or jump, a periodic component, or a combination of these (Haan 1977). The time series analysis intends at detection and quantitative description of each of the generating processes underlying a given sequence of observations (Shahin et al. 1993). There are four major steps involved in a time series modeling (McCuen 2003): (i) detection, (ii) analysis, (iii) synthesis, and (iv) verification. In the detection step, systematic components of the time series such as trends and periodicity are identified. In the analysis step, the systematic components are analyzed to identify their characteristics, including magnitudes, form and their duration over which the effects exist. In the synthesis step, information from the analysis step is accumulated to develop a time series model and to evaluate goodness-of-fit of the developed model. Finally, in the verification step, the developed time series model is evaluated using independent sets of data. For further details of the time series analysis, the reader is referred to the specialized books on time series analysis, such as Yevjevich (1972), Salas et al. (1980), Bras and Rodriguez-Iturbe (1985), Cryer (1986), Clarke (1998), and Machiwal and Jha (2012). Most statistical analyses using hydrologic time series are based on fundamental assumptions of the time series characteristics, which include; the series is homogenous and follows normal probability distribution, stationary, free from trends and shifts, non-periodic with no persistence (Adeloye and Montaseri 2002).

About five decades ago, hydrologic application of time series modeling was confined up to surface water problems, especially for analyzing the hydrologic extremes, such as floods and droughts (McCuen 2003). However, with enlarging domain of statistical hydrology over the past few decades, time series analyses presently encompass the problems of surface water as well as groundwater systems (Shahin et al. 1993; Machiwal and Jha 2006). With such a broad domain, time series analysis has emerged as a powerful tool for analyzing surface and subsurface hydrologic time series data. An extensive review on the applications of time series analysis in surface water hydrology, climatology and groundwater hydrology has been presented by Machiwal and Jha (2006). That review revealed that although several studies deal with the application of time series analysis in surface water hydrology, the application of time series analysis in groundwater hydrology is highly limited. Salient studies analyzing characteristics of the hydrogeochemistry time series are reviewed in the following sub-sections.

Normality of groundwater quality variables

The assumption of presence of normality in time series of groundwater quality variables is very crucial in obtaining reliable results of the parametric statistical tests (USEPA 1996). In the past studies dealing with geochemistry data, normality tests are applied to the point data of the groundwater quality variables, and not to the spatial data/maps on GIS platform. For example, Mouser et al. (2005) tested pH, electrical conductivity (EC) and calcium concentration data from the Molly Bog peatland located between Stowe and Morristown in Vermont, USA, for presence of normality using normal probability plot and Shapiro–Wilk test. Results suggested that calcium concentration did not follow a normal distribution, which was then subjected to logarithmic-transformation. Chou (2006) suggested use of normal probability plot for testing normality in environmental data, and transforming the non-normal data by applying logarithmic or Box–Cox transformations (Box and Cox 1964). Aguilar et al. (2007) applied three tests (i.e., Shapiro–Wilk, Shapiro–Francia, and D’Agostino tests) to identify presence of normality in nitrate concentrations of 24 sites located in Hesbaye chalk aquifer of the Geer basin of Belgium. The Shapiro–Wilk test was employed for a sample size < 50 , and the Shapiro–Francia test for a sample size ≥ 50 . In another study, Kolmogorov–Smirnov test was used to examine normality of chemical (i.e., ammonium, nitrate, nitrite, soluble reactive phosphorus and total phosphorus) and microbiological (i.e., bacterial abundance, cell biomass and bacterial biomass) variables of groundwater samples collected in Doñana aquifer system of southwest Spain (Ayuso et al. 2009). The non-normal variables identified by the Kolmogorov–Smirnov test were then transformed to make them normal. Nas and Berkay

(2010) used frequency plot and quantile–quantile plot to check normality of urban groundwater quality data (i.e., pH, EC, chloride, sulfate, hardness and nitrate) for 177 sites in Konya City (Turkey). Results revealed non-normality of EC, chloride, sulfate, hardness and nitrate, which were then normalized by log-transformation. Hosseini and Mahjouri (2014) employed Anderson–Darling test to assess normality of nitrate concentrations in Karaj aquifer of Iran. The non-normal data were transformed by using logarithmic transformations. Jovein and Hosseini (2017) examined normality of EC of the groundwater samples collected from Mahvelat Plain located in the northeastern part of Iran by normal quantile–quantile and frequency plots, and transformed the data by applying Box–Cox transformation. Recently, Leite et al. (2018) evaluated multivariate normality using Shapiro–Wilk generalized test of 14 water quality parameters for 12 sites comprising three distinct micro-regions of Santa Catarina State (four sites per region) in the municipality of Ponte Alta do Norte and São Cristóvão do Sul, Brunópolis, and Curitiba located in Marombas River basin of southern Brazil.

In addition to statistical tests, skewness and kurtosis values were also computed to test normality of trace elements (i.e., arsenic, lead, cadmium, and aluminum) present in groundwater of Dhemaji district of Assam, India (Buragohain et al. 2010). All trace elements followed non-uniform distribution in the area. Similarly, skewness and kurtosis values besides Kolmogorov–Smirnov test were used to evaluate normality of groundwater quality parameters in Amol-Babol Plain (Narany et al. 2014) and Torbat-Zaveh Plain, Khorasan Razavi (Nematollahi et al. 2016) of Iran, Modena Plain of central Italy (Barca and Passarella 2008), Antonio-El Triunfo mining district, Baja California Sur of Mexico (Wurl et al. 2014) and Pingtung Plain of Taiwan (Jang et al. 2016). The normality of trace elements present in groundwater of Greece was evaluated by computing skewness and kurtosis values, and Box–Cox transformations were used to normalize data (Dokou et al. 2015). Using Shapiro–Wilk test, Noshadi and Ghafourian (2016) checked normality of groundwater quality parameters (i.e., calcium, chloride, bicarbonate, magnesium, sodium, nitrite, nitrate, pH, sulfate, total dissolved solids (TDS), hardness, EC, and sodium adsorption ratio) at 298 sites in Fars province of Iran. Results indicated presence of normality in all the parameters except magnesium that was later on considered normal looking at value of skewness (< 2). In addition, Gan et al. (2018) log-transformed chloride, sulfate and arsenic with substantial skewness and kurtosis values to improve the normality of distribution in groundwater of Jiangnan Plain, located in central Yangtze River Basin of central China.

It is clear from the above discussion that a variety of techniques has been used in literature to examine presence of normality in the time series of groundwater quality variables.

The normality of groundwater quality variables is mainly evaluated by applying graphical (histograms, box–whisker plots, normal probability plots, quantile–quantile plots, etc.) and statistical tests (skewness and kurtosis, Chi-square (χ^2) test, Kolmogorov–Smirnov test, Lilliefors test, Anderson–Darling test, Cramér–von Mises test, Shapiro–Wilk test, probability plot correlation coefficient, Jarque–Bera test, D’Agostino Pearson Omnibus test, etc.). Furthermore, normality is tested for point data of individual sites without their GIS integration. In fact, a methodology for examining normality of the spatially distributed values of the point estimates is currently lacking that restricts the researchers to test normality as well as to present the normality test results in a spatially distributed manner.

Trends in groundwater quality variables

Identifying trends and understanding their underlying mechanisms can help make appropriate decisions for groundwater quality management (McBride 2005). Loftis (1996), for the first time, presented a review on trends in groundwater quality by discussing case studies from different parts of the world ranging from regional to local scales. Based on the review, the exact meaning of “trend” was emphasized as a critical step for groundwater quality studies in both temporal and spatial contexts. In literature, studies dealing with trend identification in groundwater quality are very less in comparison to those exploring trends in surface water quality (Taylor and Loftis 1989; Loftis 1996). Since the last decade, application of statistical trend tests in groundwater quality studies has been receiving an increasing attention (Visser et al. 2009; Kaown et al. 2012; Machiwal and Jha 2015; Yazdanpanah 2016; Koh et al. 2017).

Since the inception of the GIS technology in groundwater studies during 1990s, in studies involving trend identification of groundwater quality variables, researchers adopted the GIS mainly to present point-wise results of trend tests geographically for depicting spatial distribution of presence/absence of the trends (e.g., Mendizabal et al. 2012). Review of the literature revealed that most of the trend detection studies have dealt with individual parameters of the groundwater quality, such as sulfate (Malapati et al. 2011), chloride (Scanlon et al. 2010), hardness (Hudak 2001), etc. Few researchers explored trends in multiple groundwater quality parameters (e.g., Machiwal and Jha 2015; Masoud et al. 2016). Machiwal and Jha (2015) detected trends in 15 groundwater quality parameters of 53 sites located in Udaipur district of Rajasthan, India, by applying three tests, i.e., Kendall’s rank correlation test, Spearman rank order correlation test and Mann–Kendall test. The presence of serial correlation in all groundwater quality parameters was also tested before applying the Mann–Kendall test as the outcome of this test gets affected under the presence of serial

correlation (Yue et al. 2002). Similarly, trends in multiple groundwater quality parameters (23 quality variables) were assessed for 20 sites located in Tanta district of Egypt by applying Mann–Kendall test (Masoud et al. 2016). Results indicated that the statistically significant trends at 5% significance level were remarkable for the total hardness, total alkalinity, TDS, iron, manganese, nitrite, ammonium, phosphate and silica. However, the major focus of the past studies has been on identifying trends in nitrate (Hudak 2000a, b; Scanlon et al. 2008; Bronson et al. 2009; Enwright and Hudak 2009; Chaudhuri et al. 2012). Since 2000, researchers have been investigating how best to regionalize the nitrate concentration trends in the groundwater.

Statistical trends are generally detected by two approaches: parametric and nonparametric (Machiwal and Jha 2015). The most widely used parametric method is regression test, which is more powerful than the commonly employed nonparametric Mann–Kendall test, but the former approach requires the data be independent and normally distributed (Gilbert 1987; Bethea and Rhinehart 1991), and the latter approach is free from such normality assumption. Helsel and Frans (2006) developed a regional-Kendall method based on the principle of the seasonal-Kendall test to determine regional trends in the groundwater quality. The performance of the regional-Kendall test was found satisfactory by researchers such as Frans (2008), Sprague and Lorenz (2009), and Kaown et al. (2012). Recently, Lopez et al. (2015) developed a methodology for application of the regional-Kendall test in GIS platform by using geostatistics. In addition, Yazdanpanah (2016) made an effort to integrate results of linear trend analysis with GIS using geostatistical modeling of slope values of the linear trend model.

Similar to the normality-testing of the point-wise groundwater quality variables, it is evident that the literature studies have identified trends in a variety of groundwater quality parameters over individual sites in different parts of the world. However, trend detection studies using spatially distributed GIS maps of the groundwater quality variables are not found in the literature. The major cause for non-consideration of spatial maps for trend assessment is non-availability of a methodology for coupling statistical trend tests with GIS to identify the temporal trends directly into the GIS framework. In contrast to normality-testing, a large number of the studies used GIS to present the results of the trend tests for groundwater quality time series over space. Recently, Kumar et al. (2017) developed a standard methodology to identify spatial trends using spatial raster datasets in GIS framework by coupling three statistical tests (i.e., Kendall rank correlation test, Spearman rank order correlation test, and Mann–Kendall test) with GIS. However, the methodology developed was demonstrated through a case study to identify trends in rainfall of Gujarat state, India using satellite datasets. There is a need to employ such methodology

for trend identification in time series of groundwater quality variables.

Persistence

Persistence of a time series is the tendency of the successive data to “remember” their antecedent data and to be influenced by them (Giles and Flocas 1984). In other words, it is defined as the correlational dependency of order or time lag “ k ” between each i th element and the $(i-k)$ th element of time series (Kendall 1973), and is measured by autocorrelation (i.e., correlation between two terms of the same time series). In hydrogeochemistry, persistence testing is reported in a few studies (e.g., Jones and Smart 2005). They investigated internal structure of long-term nitrate concentration records for five karst springs in Mendip Hills, England (UK) by using stochastic autoregressive modeling. The results indicated the significant short-term persistence of 1–2 months in three of five springs.

It is also seen that except normality and trends, the remaining characteristics of the time series such as homogeneity, stationarity, periodicity, and persistence, are generally ignored in the studies dealing with groundwater quality over spatial and temporal scales.

Evaluation of groundwater quality data using GIS-integrated multivariate statistical methods

Application of multivariate statistical methods in groundwater quality studies

As it was put forth by J.D. Hem in his seminal work (Hem 1985), water chemistry (i.e., hydrochemistry)—the field of study mainly concerned with chemical and physical properties of natural waters—“hardly qualified as a scientific discipline” until the late 1950s. Interest to groundwater chemistry (i.e., hydrogeochemistry) studies occurred even much later and was not extensive until 1972 (Niu et al. 2014). Furthermore, at the time being, water quality testing was mostly an expensive and arduous endeavor requiring a variety of volumetric, gravimetric, colorimetric, turbidimetric, complexometric, and potentiometric procedures (Rainwater and Thatcher 1960). Today, with a wide array of modern analytical instruments and technologies at our disposal (APHA-AWWA-WEF 2017), it is possible to identify and quantify a great number of chemical constituents (inorganic and organic) and to measure various physical parameters in water samples of different matrix complexity (ranging from freshwater to brine) with a better measurement accuracy/precision and at a lesser cost than ever before. Obviously, statistical analysis of such multidimensional datasets, acquired at

different spatial and temporal scales, require computationally efficient and sophisticated techniques to close the widening gap between our data-generating and data-analyzing capabilities. Especially after 1980s, dramatic increase in the processing and storage capacities of computer hardware, coupled with the emergence of powerful multi-tasking (GIS-based) software packages integrating relational database management (RDBM), statistical/spatial analysis, and two or three dimensional (2-D/3-D) visualization tools have made possible to analyze such large datasets for extracting and summarizing relevant quantitative information hidden in data. This type of insight cannot be solely gained from the conventional graphical plots (Hem 1985; Zaporozec 1972; Güler et al. 2002; Machiwal and Jha 2010) that are used in visualization and interpretation of water quality data.

The chemical composition of natural water is derived from many different complex processes and sources (natural and/or anthropogenic), all of which imprint a unique physicochemical signature on the water constantly recycling through the Earth's spheres (atmosphere, hydrosphere, geosphere, and biosphere). Therefore, statistical analysis of hydrochemical data entails the simultaneous evaluation of all the chemical and physical parameters (or variables) measured, since water quality is a function of these properties (Williams 1982). One of the ways to accomplish such an evaluation is through multi-variate statistical analysis (MVSA). Generally speaking, the main objective of the MVSA is to simplify the data matrix (composed of p variables and n cases) by finding associations among dataset variables (called R-mode analysis) and/or cases (called Q-mode analysis) (Dalton and Upchurch 1978). Ideally, the information extracted from the data matrix should be easily understood and potentially useful, providing new insights about data. More often, MVSA is a stepwise procedure with the obvious first step getting acquainted with data at hand. Therefore, prior to conducting any formal statistical analysis (i.e., univariate, bivariate, and multivariate), all variables in the dataset should be carefully scrutinized for data quality, measurement/entry errors, missing/censored values, and outliers in order to verify data integrity/consistency and identify the variables/cases violating certain rules and/or method assumptions (Güler et al. 2002). Nonetheless, most of these data quality issues can be minimized, if not resolved completely, by developing and implementing reliable quality assurance/quality control (QA/QC) protocols for collection, handling, and analysis of water samples, both in the field and in laboratory. While most quantitative analytical data, especially the ones related to naturally occurring trace elements, do not lend themselves directly to inferential MVSA due to problems related to non-normality and heteroscedasticity (i.e., heterogeneity of variances), there are techniques available (e.g., Box–Cox transformations (Box and Cox 1964; Box et al. 1978) and z score scaling, i.e., standardization)

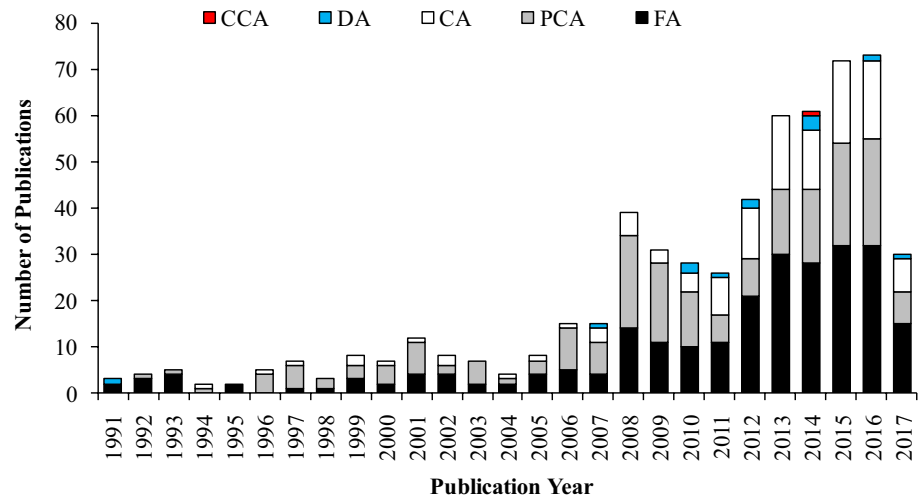
(Johnson and Wichern 1992) to improve the overall data distribution for elucidating latent associations among data variables and/or cases. Although such statistical associations do not directly establish cause-and-effect relationships, they can assist in creating hypotheses to make viable predictions about the underlying complex processes and phenomena responsible for the data variance and noise (Güler et al. 2002, 2017). However, the challenge is to decide which MVSA methods are best suited for the problem at hand, and understanding their theoretical assumptions and inherent strengths/weaknesses.

Up until late 1960s, statistical treatment of the water quality variables was mainly limited to univariate and bivariate numerical analysis (e.g., calculation of ionic ratios, mean/extreme values, and correlation coefficients) and graphical displays (e.g., frequency distributions and scatter plots) (Hem 1970). Most of the MVSA methods in common use today for tackling Earth Science problems have been adopted from other scientific disciplines (e.g., physics, astronomy, biology, and social/behavioral sciences), where the use of these methods was widespread long before the computer era, thanks to Hollerith's electromechanical punch-card tabulator and dedicated human computers. Pioneering applications of the MVSA methods in water-related fields occurred much later, and was not as extensive until 1990s. A simple bibliometric analysis (this study) of peer-reviewed journals (published from 1980 forward) listed by The Institute for Scientific Information (ISI) Web of Science online database revealed that factor analysis (FA) is by far the most frequently used MVSA method in groundwater studies, followed by principal component analysis (PCA), cluster analysis (CA), and discriminant analysis (DA) (Fig. 2). However, applications of canonical correlation analysis (CCA) and Correspondence Analysis to groundwater studies are extremely rare or non-existent. It is further evident that the number of studies involving use of the MVSA techniques in hydrogeology has an overall increasing trend over the years, with a significant increment during the period 2011–2016. In the following sections, we briefly introduce salient exploratory MVSA methods, along with examples from the scientific literature that focus on their different applications, for extraction of relevant information concealed in data.

Eigenvector methods: factor analysis vs. principal component analysis

FA and PCA can be described as “eigenvector” methods for finding lines and planes of closest fit to systems of points in multidimensional space, mathematical foundation of which was first established by Pearson as early as 1901 (Thurstone 1931). Both FA and PCA have been extensively applied in many disciplines (especially in social and behavioral sciences) mainly for data reduction purposes, (e.g., reduction

Fig. 2 Bar charts depicting usage and growth of multivariate statistical analysis methods in SCI-expanded publications related to hydrogeochemical studies (as of July 4, 2017). CCA canonical correlation analysis, DA discriminant analysis, CA cluster analysis, PCA principal component analysis, FA factor analysis



of variables or cases). The data reduction using these methods is achieved through finding the directions of maximum variances (i.e., eigenvalues) in a multivariate dataset and representing them in a much lower dimension (usually 2–5) than the original dataset. In both methods, the variance analysis involves decomposition of the matrix of correlations, which presents interrelations among all pairs of the original variables. However, PCA is often preferred as a method for data reduction, while FA is often used when the goal of the analysis is to detect the structure (i.e., a few underlying, but unobservable, latent constructs or factors) in a dataset (Suk and Lee 1999). Indeed, combining two or more correlated variables (or vectors) into one “factor” or “principal component (PC)”, exemplifies the basic idea of FA and PCA. The new factors and PCs extracted by FA and PCA (respectively) are uncorrelated and ordered so that each successively extracted factor (e.g., F1, F2, and so on) or PC (e.g., PC1, PC2, and so on) accounts for a lesser amount of variance of the original dataset than the previous one (Davis 1986; Brown 1998). FA and PCA are occasionally mistaken as the same MVSA method, probably because of the apparent similarities in the terminology and methodology used for both. Despite the similarities in the terminology, there are distinct differences between the methodologies of FA and PCA methods. However, in most cases, FA and PCA usually yield very similar results, if communalities (i.e., proportion of variance that each variable or case has in common with other variables or cases) are close to unity.

Assumptions of both FA and PCA include that: (i) each original variable follows normal distribution; (ii) original variables display linear relationships; (iii) there are no outliers in data; (iv) sample size is adequate ($n \geq 50$, or $n \geq 100$ for more stable estimates) and balanced (i.e., case to variable ratio is at least 5). These multi-step MVSA methods (FA and PCA) have been used extensively in hydrogeochemical studies (both in R- and Q-modes) to extract the relevant

information hidden in data matrices, e.g., to: (i) extract and ordinate the most important and influential parameters (i.e., physical and chemical variables) responsible for the spatial and/or temporal variations in water quality (Ashley and Lloyd 1978; Melloul and Collin 1992; Ribeiro and Macedo 1995; Reghunath et al. 2002; Thyne et al. 2004; Cloutier et al. 2008); (ii) ascertain the similarities/dissimilarities or continuity/overlap in spatially and/or temporally distributed groundwater samples (i.e., cases) (Güler et al. 2002; Dalton and Upchurch 1978; Usunoff and Guzman–Guzman 1989; Farnham et al. 2002); (iii) reveal underlying latent factors (e.g., key processes, phenomena, sources, and end-members) that account for the structure of the hydrochemical data (Dawdy and Feth 1967; Melloul and Collin 1992; Suk and Lee 1999; Meng and Maynard 2001; Lambrakis et al. 2004; Güler et al. 2017; Kazakis et al. 2017; Busico et al. 2018), and (iv) produce data for further investigation or for other methods (e.g., factor score mapping, multiple regression, cluster analysis, GIS analysis, etc.) (Dalton and Upchurch 1978; Subbarao et al. 1996; Suk and Lee 1999; Güler et al. 2012).

Many past studies applied FA or PCA in combination with GIS techniques to identify the anthropogenic and natural hydrogeologic processes functioning in the aquifer systems [e.g., (Thyne et al. 2004; Dragon 2006; Güler et al. 2012; Petrişor et al. 2012)]. However, precise and proper spatial analyses integrating PC scores with GIS-based geostatistical modeling are rarely carried out (e.g., Güler et al. 2012; Narany et al. 2014; Machiwal and Jha 2015).

Cluster analysis

The term “Cluster Analysis”, first introduced by Tryon (Tryon 1939), encompasses a wide variety of classification algorithms applied in many fields (including hydrogeochemistry) to organize data variables and/or cases into

homogenous and non-overlapping subsets or groups, called clusters (Hartigan 1975). Using this method, the original data matrix, composed of p variables and n cases, is partitioned into k number of subsets (where k is generally much smaller than p (in R-mode) and/or n (in Q-mode)); hence, data reduction is achieved. In general, the members of each cluster share similar characteristics (e.g., in terms of chemical composition) compared to non-members (members belonging to other clusters). In this method, grouping of individual variables and/or cases is generally achieved through an iterative process, where the number of clusters (k) may or may not be known a priori. The cluster centroids (or means) obtained from the resulting partition can be used as representative members (a.k.a. prototypes) of their respective groups. In hydrochemical studies, commonly used partitioning algorithms include hierarchical clustering (joining or tree clustering) and K-means clustering.

The hierarchical clustering analysis (HCA) employs various types of distance (similarity/dissimilarity) measures and linkage methods (i.e., amalgamation rules) (Sneath and Sokal 1973; Hartigan 1975) and the choice of which combination to use does not have an easy answer and greatly affects outcomes (Güler et al. 2002). One of the most widely used combination in HCA is the Euclidean distance (as distance measure) and Ward's method (for linkage), which forms distinct and easily interpretable clusters that may be significant in the hydrochemical, hydrologic, and geologic contexts (Gong and Richman 1995; Güler et al. 2002). Since, the results of HCA are mostly presented in a tree-like 2-D diagram called dendrogram (Davis 1986), the method is generally appropriate for partitioning small datasets (Güler et al. 2002). However, this apparent shortcoming can be overcome by employing a multi-step clustering approach (e.g., pre-clustering and then re-clustering) (Güler and Thyne 2004a) or using another MVSA method (e.g., PCA) first for data reduction and noise filtering (Pirkle et al. 1984). Dendrograms can also be spatially projected in 3-D, in a map form using color-coded clusters (Forina et al. 2002), but to our knowledge this technique has not been used in hydrochemical and hydrogeochemical studies. The HCA is generally accepted as a semi-objective procedure, not requiring a priori specification of the number of clusters (k), where their numbers is usually defined post-process (somewhat subjectively), drawing a line (i.e., phenon line) that cuts through dendrogram branches at a certain distance value (Güler et al. 2002). Unlike HCA, K-means clustering (KMC) is a nonhierarchical method that can allow classification of a substantially large number of samples (Gong and Richman 1995; Pacheco 1998). KMC follows a simple iterative procedure, which assumes exactly k number of random cluster centers at the onset of the analysis (MacQueen 1967). If k is unknown a priori, a subjective bias may be introduced into the results. The KMC tries to define exactly k different cluster centroids

(one for each cluster) with the greatest possible distinction. From the computational point of view, KMC can be thought of as analysis of variance (ANOVA) in reverse (Güler et al. 2002). The KMC process is initiated with k random clusters, and then objects to be clustered are iteratively relocated between those clusters with the aim to: (i) reduce within-cluster variance and (ii) increase between-cluster variances (Pacheco 1998). However, the resulting KMC partition is highly sensitive to the initial randomly selected cluster centers. Executing multiple KMC runs on the dataset can help to minimize this effect. The results from KMC are typically presented in matrix form, which shows members of each cluster and their distances from respective cluster centers (Güler et al. 2002).

CA techniques mentioned here relies on assumptions such as normal distribution and equal variance (homoscedasticity) of the water chemistry data variables that are continuous in nature (Alther 1979). In addition, the use of variables having specific relationships or displaying a high intercorrelation among themselves (i.e., multicollinearity) may cause unwanted redundancies in the clustering process (Güler et al. 2002). However, CA using factor scores obtained from the factor analysis (FA) can be used in order to reduce multicollinearity (Suk and Lee 1999). Outliers should also be treated with caution, since they tend to strongly distort the results. Another important and essentially unsettled issue in cluster analysis is the "cluster validity problem" (Hardy 1996), which mainly involves determination of the "true" number of groups (k) in a dataset (mostly unknown a priori). In hydrogeochemical studies, the spatial coherence of the statistically defined groups (e.g., similarity/proximity of geographical locations, altitudes, and distances of within and between group members) can be verified using GIS-based spatial analysis techniques for cluster validation purposes, which may also provide insights into aquifer heterogeneity/connectivity and the processes governing water quality (Güler and Thyne 2004a). As a general rule of thumb, distinctly different hydrogeochemical groups should be both statistically and spatially (in a geographical sense) well separated, due to increasing water-rock interactions along hydrological flow-paths (Thyne et al. 2004).

Since the late 1970s, CA has been successfully applied to water-chemistry data in many groundwater studies to: (i) classify samples into distinct hydrogeochemical groups (Ashley and Lloyd 1978; Riley et al. 1990; Johnson and Wichern 1992; Suk and Lee 1999), (ii) identify hydraulic connections between surface and deep zones (Williams 1982), (iii) interpret groundwater flow (Ochsenkühn et al. 1997), (iv) find optimal number of natural clusters (Pacheco 1998), (v) classify samples from different aquifers (Steinhorst and Williams 1985; Saleh et al. 1999), and (vi) evaluate temporal changes in groundwater composition (Ribeiro and Macedo 1995; Suk and Lee 1999; Berzas et al. 2000).

A number of researchers (Farnham et al. 2000; Meng and Maynard 2001; Güler et al. 2002; Güler and Thyne 2004a; Thyne et al. 2004; Helstrup et al. 2007; Cloutier et al. 2008) used R- and Q-mode cluster analyses, in conjunction with other MVSA, geochemical (modeling), and spatial analysis techniques (e.g., GIS) for hydrogeological and hydrogeochemical site characterization in groundwater studies with scales ranging from local to regional.

Discriminant analysis

The main purpose of discriminant (function) analysis (DA) is to determine a set of characteristics (i.e., variables, p) that permit for the best prediction (discrimination) between two or more naturally occurring a priori defined groups ($k \geq 2$) within the dataset or assigning new objects (i.e., cases, n) accurately into these homogeneous groups on this basis (Izenman 2013). The basic notion underlying DA is to decide whether groups differ with regard to the mean of a predictor variable, and then to use that variable to predict group membership of new cases. The prediction is achieved by linear discriminant functions (Johnson and Wichern 1992; Wunderlin et al. 2001), which are vectors (linear combination of the selected independent variables) in the directions of optimal separation between the groups. DA initially requires a reference set of (representative) samples for each group, for “training” purposes. DA is generally a stepwise procedure, with forward- and backward-modes, where variables are added or removed one-by-one in a sequential manner to improve the separation between groups (Machiwal and Jha 2010). DA algorithm tries to maximize between group variance–covariance and minimize within-group variance–covariance under simultaneous consideration of all analyzed features. The impact of each variable on the discriminant function can be assessed by comparing their partial F values, where the higher the value is the more impact it has on the discriminant function. In addition, the Wilk’s lambda (λ) value is used as a measure of the statistical significance of the discriminatory power of the model ($\lambda = 0$, perfect discriminatory power and $\lambda = 1$, no discriminatory power), while Mahalanobis distance statistic (D^2) is used to assess for separation of groups. Computationally, DA is analogous to the one-way/multivariate analysis of variance methods (ANOVA/MANOVA). When three or more groups (k) present, the method is referred to as multiple discriminant analysis (MDA), which has close associations with other MVSA methods, including multiple regression analysis, FA, and canonical correlation. DA or MDA can be used to classify and, thus, to confirm the groups found by means of CA. DA relies on the same assumptions that are required for the CA (e.g., normal distribution, homogeneity of variances/covariances, multicollinearity, and no outliers) (Izenman 2013). In addition, the variables that are used

to discriminate between groups should not be completely redundant with the other variables. In other words, if a variable (e.g., total dissolved solids) is the sum of a number of other variables (e.g., ionic constituents) that are also being evaluated, then the “ill-conditioned matrix” problem may occur. DA can be used successfully when the dependent variable is “categorical” and the independent variables are “metric” and normally distributed.

In hydrogeochemical studies, DA is infrequently used and has been applied for the assessment of spatio-temporal variations in datasets (Steinhorst and Williams 1985; Wunderlin et al. 2001), where site (spatial) and season (temporal) can be coded as grouping variables, while the measured physical and/or chemical parameters constitute independent variables (Machiwal and Jha 2010). The further details on the procedure can be found in Cooley and Lohnes (1971) and Johnson and Wichern (1992).

Canonical correlation analysis

Canonical correlation analysis (CCA) is considered as one of the correlation techniques (Hotelling 1936). However, it is different from the FA or PCA in spite of certain similarities in concept and terminology. In general, it is used to investigate the intercorrelation between two datasets of variables, whereas FA, PCA or empirical orthogonal functions detect a pattern of relationship within one dataset (Clark 1975). The CCA can be used for examining presence of any similar kind of pattern that may occur simultaneously in two different datasets, and if it is present, then the correlation between associated patterns is calculated. In hydrogeochemistry studies, the application of CCA could not be found in literature. However, for measuring trophic status of reservoirs and lakes, Cairns et al. (1997) applied CCA on the water parameters (namely chlorophyll, total suspended and dissolved solids, and turbidity) and digital values of three bands and numerous band ratios of SPOT (Système Pour l’Observation de la Terre) satellite data. The results indicated that the turbidity and chlorophyll contributed 0.91 and 0.76, respectively, to the first canonical water variable showing a good relationship. It is worth mentioning that the DA is one of the special cases of CCA.

Application of geostatistical modeling in groundwater quality evaluation

Geostatistical modeling techniques were originally developed and applied in geological studies for estimating mineral concentrations in ore bodies and recoverable reserves (David 1977; Journel 1974; Journel and Huijbregts 1978). It is seen from the literature that the first attempts of geostatistical-modeling application to geochemical data were

made by David and Dagbert (1975) and David (1977). During the 1940s, an important contribution of geostatistics in meteorology was made by the Soviet School of Meteorology (Drozdov and Shepelevskii 1946). Later on, Gandin (1965) and Kagan (1967) emphasized the need of recognizing spatial variability along with quantification of estimation error. Kriging is the widely used geostatistical technique developed by Matheron (1965, 1973). In hydrogeology, Delhomme (1978) paved the way for the geostatistical-modeling applications. Application of the geostatistical-modeling techniques in groundwater quality studies was very limited up to the end of 1990s. However, with integration of GIS, use of geostatistical-modeling techniques in groundwater quality evaluation significantly increased. After 2000, studies involving GIS-integrated geostatistical-modeling techniques mushroomed in literature. Cooper and Istok (1988a, b) made an excellent effort by developing a comprehensive methodology for applying geostatistics to the problems of groundwater contamination, and demonstrated its application through a case study at the Chem-Dyne Superfund site in Ohio, USA. Istok and Cooper (1988) developed techniques for combining local estimates obtained by kriging to obtain global estimates and estimation errors for the expected contaminant concentration in any specified portion of the contaminant plume. An overview of the basic concepts of the geostatistics and its proposed linear and nonlinear kriging estimation techniques is provided by the ASCE Task Committee (1990a). The ASCE Task Committee (1990b) reviewed the applications of the geostatistical-modeling techniques in groundwater hydrology under the five major sections: (i) mapping, (ii) simulation of hydrological variables, (iii) estimation using the flow equations, (iv) sampling design, and (v) geostatistical-modeling applications in groundwater system management.

In 1980s, few researchers applied geostatistics in hydrogeochemistry studies (e.g., Myers et al. 1982). Myers et al. (1982) developed four variogram models (i.e., linear, constant-linear, concave and convex) for 12 variables (U, B, Ba, Ca, Li, Mg, Mo, As, V, SO₄, Specific conductance, and total alkalinity) in Ogallala formation and Permian geologic units in Texas, USA. The results of the geostatistical modeling were compared with the inverse distance weighting (IDW) technique, which revealed a better performance of the IDW technique in spatial interpolations of the variables. Rouhani and Hall (1988) used geostatistical techniques for the design of a regional shallow groundwater quality monitoring network in the Dougherty Plain, located in southwest Georgia, USA. In Spain, reliability of groundwater-quality monitoring network in controlling saltwater intrusion was assessed by using lognormal kriging for mapping chloride distribution in the Llobregat delta confined aquifer of Barcelona (Candela et al. 1988). Bárdossy and Kundzewicz (1990) applied two geostatistical methods (i.e., point kriging and

intrinsic random function of order 'k') for detection of outliers in chloride and total hardness data of groundwater from the Upper Rhine Valley, extending across three countries: France, Germany, and Switzerland. Bjerg and Christensen (1992) evaluated horizontal variations in groundwater-quality parameters (i.e., pH, alkalinity, Cl, NO₃, Ca and K) in a shallow sandy aquifer located in western part of Denmark. Results indicated substantial variations in all parameters even at smaller distances. Istok et al. (1993) presented a case history of the alluvial aquifer underlying the Malheur River Basin, Oregon, USA, where isotropic and spherical geostatistical models were applied for estimating pesticide concentrations from the measured nitrate and pesticide concentrations under the limited sampling of the pesticides. Rautman and Istok (1996) presented a geostatistical framework for probabilistic assessment of groundwater contamination and illustrated the approach using synthetic data of a hypothetical site. The approach is further demonstrated through a case study in agricultural area in the Lower Malheur River Basin and the Western Snake River Plain near the eastern Oregon, USA (Istok and Rautman 1996). Pebesma and de Kwaadsteniet (1997) prepared spatial maps of 25 groundwater-quality variables based on median measurements of 425 sites in 4 × 4 km block in the Netherlands using block kriging. Their study quantified the effect of monitoring network density, and evaluated changes in the groundwater quality over a span of 20 years.

Ordinary kriging and cokriging were compared for studying spatial distribution of nitrate in Lucca Plain aquifer of Central Italy (D'Agostino et al. 1998). Results indicated that the cokriging improved the estimation and reduced the uncertainty in terms of estimation variance. In multivariate geostatistical problems, two related variables are used to improve estimation of the primary variable by using the secondary variable. The ordinary kriging and cokriging have a smoothing effect causing underestimation (or overestimation) of the variable due to large (or small) sample values in cross-validation. This smoothing effect was reduced by applying Gaussian random-process based principle in simulating kriged and cokriged estimates using chloride (primary variable) and resistivity (secondary variable) data in Horonobe area of northern Japan (Lu et al. 2016). Sánchez-Martos et al. (2001) first applied the PCA technique to identify three factors (i.e., sulfate, thermal and marine influences) that affect groundwater processes in the detrital aquifer of the Bajo Andarax (Almería, Spain). Then, the identified three factors were analyzed using ordinary block kriging to obtain their spatial distribution. Geostatistical approach has been used with Bayesian analysis for contaminant source identification by developing a methodology to estimate release history of a conservative solute (Snodgrass and Kitanidis 1997). This approach is subsequently extended to the estimation of the antecedent

distribution of a contaminant at a given point back in time (Michalak and Kitanidis 2004). Later on, Sun (2007) further extended this approach to develop a robust geostatistical approach to contaminant source identification by solving the linear estimation problems. Empirical Bayesian kriging is another technique of kriging family, which is different from other kriging methods as the former uses an intrinsic random function for spatial interpolation (Gupta et al. 2017). This technique is rarely used for spatial interpolation of groundwater quality variables (e.g., Mirzaei and Sakizadeh 2016).

After the year 2000, several studies have applied geostatistical-modeling techniques for mapping spatial variability of the chemical concentrations for groundwater quality assessment/evaluation (e.g., Goovaerts 1999; Castrignanò et al. 2000; Yu et al. 2003; Mouser et al. 2005; Schaefer and Mayor 2007; Machiwal and Jha 2015). At present, abundant studies on this aspect exist in literature, and there has been an increasing trend in appearance of such studies in research journals after the year 2000. Salient studies, reported after 2000, dealing with the use of geostatistical-modeling techniques for the groundwater quality assessment/evaluation are enlisted in Table 1. It is revealed that ordinary kriging is the most widely used geostatistical-modeling technique in groundwater quality studies. The spatial distribution of almost all kind of groundwater-quality parameters along with scores of the principal components obtained through PCA is determined in different parts of the world. It is also observed that the estimation error of the geostatistical modeling is computed in a large number of studies by using cross-validation criteria. However, some of these studies ignored the important step of validation while applying the geostatistical modeling.

Application of hybrid methods for groundwater vulnerability assessment in GIS platform

Concept of groundwater vulnerability, first introduced by Margat (1968), is based on the assumption that the physical environment may provide some degree of protection to groundwater against human activities. The groundwater vulnerability is classified into two types: specific vulnerability and intrinsic vulnerability (National Research Council 1993). Intrinsic vulnerability of an aquifer can be defined as the ease with which a contaminant introduced onto the ground surface can reach and diffuse in groundwater (Vrba and Zaporozec 1994). On the other hand, specific vulnerability is used to define the vulnerability of groundwater to particular contaminant or a group of contaminants by taking into account the physicochemical properties of the contaminants and their relationships (Gogu and Dassargues 2000). Initially, groundwater vulnerability of an aquifer is

mapped, and hence, it can be used as an assessment tool against groundwater pollution. Since the year 1968, numerous groundwater vulnerability assessment methods have been developed and applied worldwide, which are coupled with the GIS. The vulnerability assessment methods can be classified into: (i) index-based methods, (ii) quantitative or simulation models, (iii) statistical and artificial intelligence methods, and (iv) hybrid methods that are the combination of the earlier three methods mainly integrating index-based methods with statistical and artificial intelligence methods. A tree diagram illustrating classification of methods for groundwater vulnerability assessment is presented in Fig. 3. Shirazi et al. (2012) presented a review of application of the GIS-based DRASTIC method for groundwater vulnerability assessment. Later on, Wachniew et al. (2016) summarized review of intrinsic methods of groundwater vulnerability assessment. It is learnt that earlier reviews emphasized on a particular vulnerability assessment method or only index-based methods for groundwater vulnerability. Recently, Machiwal et al. (2018) presented a comprehensive review of groundwater vulnerability highlighting current status and challenges of index-based, quantitative and statistical methods including methods for source protection. However, in this paper, we focus mainly on hybrid methods developed by combining advanced statistical and artificial intelligence techniques with index-based methods. The index-based methods are parameter weighting and rating methods, which, apart from classifying the various parameters, also introduce relative weight coefficients for each factor. Such methods are usually coupled with statistical methods, so as to overcome the subjectivity of weights and ratings of each parameter.

Rupert (2001) introduced, perhaps for the first time, a hybrid approach for groundwater vulnerability assessment by using a calibration procedure. The groundwater vulnerability map, initially developed using the DRASTIC method, was modified according to its correlation with nitrate concentrations in the Snake River Basin in USA. Similarly, Panagopoulos et al. (2006) used Spearman's ρ and Kendall's τ correlation coefficients to modify both the weights and ratings of the DRASTIC parameters. A major concern in the assessment of groundwater vulnerability to nitrates constitutes the use of qualitative parameters. Kazakis and Voudouris (2015) replaced the qualitative parameters of DRASTIC method with quantitative ones, and proposed a new method to estimate groundwater vulnerability to nitrate. Additionally, nitrate concentration was correlated with grading methods in order to determine the more suitable classes of the proposed method. The grading methods of natural breaks, equal interval, quantile and geometrical intervals were used to define the class ranges of the final vulnerability to nitrate-based index, whilst sensitivity analysis and ANOVA F test statistics were used to verify the results. Other more complex hybrid methods include integration of

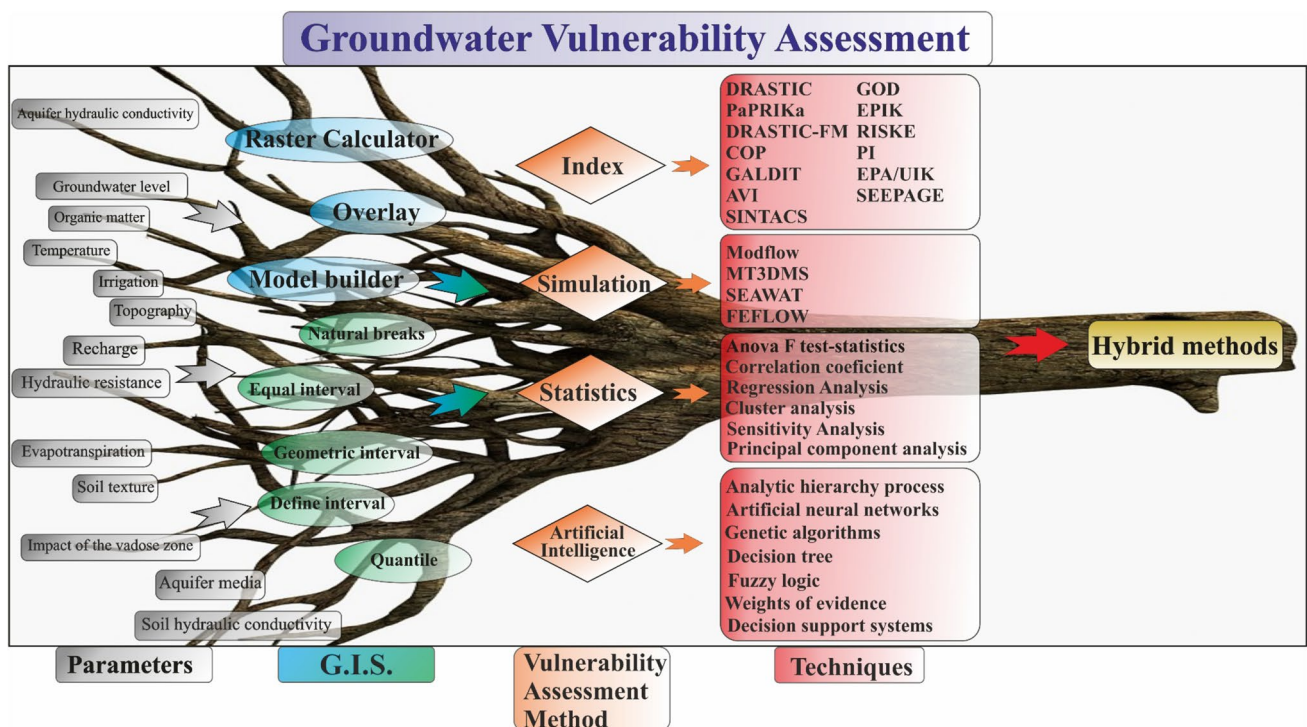
Table 1 Salient studies that utilized geostatistical techniques for mapping spatial distribution of groundwater quality after the year 2000

Country	Geostatistical-modeling technique	Validation technique	Parameters ^a	References
Australia	Ordinary kriging	Data from other study	Cl	Davies and Crosbie (2018)
Bangladesh	Cokriging	None	As	Hassan and Atkins (2007)
	Ordinary kriging	Cross-validation	Groundwater quality index, degree of contamination, heavy metal pollution index, heavy metal evaluation index, principal component scores	Bhuiyan et al. (2016)
	Ordinary kriging	Cross-validation	Groundwater quality index, degree of contamination, heavy metal pollution index, heavy metal evaluation index	Bodrud-Doza et al. (2016)
Cameroon	Ordinary and block kriging	Cross-validation	Groundwater quality index	Nshagali et al. (2015)
China	Ordinary kriging	None	TDS, Ca, Mg, Na, K, Cl, SO ₄ , HCO ₃	Chen and Feng (2013)
Egypt	Ordinary kriging	None	TDS, Fe, Mn	Masoud et al. (2018)
Germany	Ordinary kriging	None	N ₂ O	von der Heide et al. (2008)
	Ordinary and indicator kriging	Cross-validation	Cl, As, deethylatrazine	Bárdossy (2011)
Greece	Ordinary Kriging	Mean standardized prediction error	EC, NO ₃ , NH ₄ , PO ₄	Stamatis et al. (2011)
India	Ordinary kriging	Cross-validation	EC, TDS, Hardness, SAR, Mg-Ca ratio, Cl, HCO ₃ , NO ₃	Adhikary et al. (2012)
	Ordinary kriging	None	Principal component scores	Machiwal and Jha (2015)
	Empirical Bayesian Kriging	Prediction standard error	<i>F</i>	Magesh et al. (2016)
Iran	Ordinary, simple, universal, indicator, probability, and disjunctive kriging	Cross-validation	EC, SAR, Ca, Mg, Na, Cl, SO ₄ , HCO ₃	Yazdanpanah (2016)
	Ordinary kriging	Cross-validation	EC, TDS, total hardness, SAR, Na, Cl, SO ₄	Karami et al. (2018)
Italy	Disjunctive kriging	Cross-validation	NO ₃	Passarella et al. (2002)
	Disjunctive kriging	Cross-validation	NO ₃	Barca and Passarella (2008)
Japan	Ordinary kriging and cokriging	Cross-validation	Cl	Lu et al. (2016)
Korea	Ordinary kriging	None	Cl, NO ₃ , Fe, factor analysis scores	Kim et al. (2012)
	Ordinary kriging	None	Ca, Mg, Na, K, Cl, SO ₄ , HCO ₃ , NO ₃ -N	Venkatramanan et al. (2016)
Lebanon	Ordinary kriging	Cross-validation	NO ₃	Assaf and Saadeh (2009)
Malaysia	Ordinary kriging	None	Ca, Mg, Na, K, Cl, H ₄ SiO ₄ , Al, Ba, Fe, Mn, Pb, Se, Sr	Lin et al. (2012)
Saudi Arabia	Ordinary kriging	Cross-validation	Temp, EC, TDS, Salinity, Ca, Mg, Cl, NO ₃	Marko et al. (2014)
	Ordinary and probability kriging	None	TDS, pH, Ca, Mg, Na, K, Cl, SO ₄ , HCO ₃ , NO ₃	Salman et al. (2015)
Spain	Ordinary kriging	None	Principal component scores	Sánchez-Martos et al. (2001)
Taiwan	Indicator kriging	None	As	Jang et al. (2007)
	Multivariate indicator kriging	Cross-validation	Multiplication of indicator variables, average of indicator variables	Jang (2013)
Turkey	Ordinary kriging	Cross-validation	EC, pH, hardness, Cl, SO ₄ , NO ₃	Nas and Bertkay (2010)
	Ordinary kriging	Cross-validation	Fuzzy membership function values of four groundwater classes	Güler et al. (2012)

Table 1 (continued)

Country	Geostatistical-modeling technique	Validation technique	Parameters ^a	References
	Block kriging	None	Temp, DO, salt content, NO ₃ , NH ₄ , P, Cd, Co, Cr, Cu, Fe, Mn, Ni, Zn	Ağca et al. (2014)
	Ordinary kriging	None	EC, SAR, Kelly index, Mg ratio, RSC, potential salinity, %Na	Arslan (2017)
USA	Ordinary kriging	None	EC, pH, Ca	Mouser et al. (2005)
	Indicator kriging	Cross-validation	As	Goovaerts et al. (2005)

^aTDS total dissolved solids, EC electrical conductivity, SAR sodium adsorption ratio, Temp temperature, DO dissolved oxygen, RSC residual sodium carbonate

**Fig. 3** Tree diagram illustrating classification of methods for groundwater vulnerability assessment

index-based methods and artificial intelligence (AI) techniques such as fuzzy logic and artificial neural networks (ANNs).

Regression analysis has been widely used in environmental studies. In the region of Osona (NE Spain), Boy-Roura et al. (2013) used multiple linear regression and isotopes for the assessment of groundwater vulnerability to nitrates. In another study, logistic regression and weights of evidence statistical procedures were coupled with DRASTIC method for the development of two hybrid methods, which were applied in Korinthia prefecture in South Greece (Antonakos and Lambrakis 2007). In the Pearl Harbor-Honolulu aquifer

in the USA, stepwise logistic regression and capture zones of wells were coupled for the assessment of groundwater vulnerability (Mair and El-Kadi 2013). Among other techniques, Pacheco et al. (2015) applied analytic hierarchy process (AHP) for the factor weighting of DRASTIC parameters and the estimation of groundwater vulnerability in different aquifers of Portugal. Fuzzy logic and ANNs are successfully utilized to assess groundwater vulnerability (Dixon 2005a, b). Fijani et al. (2013) coupled Sugeno fuzzy logic (SFL), Mamdani fuzzy logic (MLF), ANNs, and Neuro-Fuzzy (NF) techniques in order to estimate groundwater vulnerability in the Maragheh-Bonab basin of Iran. Likewise, Larsen

fuzzy logic (LFL) was applied in conjunction with SFL and MFL for the assessment of groundwater vulnerability in the Varzeqan Plain, in northwestern Iran (Nadiri et al. 2017). In addition to above techniques, multivariate statistical analysis, e.g., PCA and CA have also been used in groundwater vulnerability assessment studies. The CA has been used to determine the most influential chemical factors in determining aquifer vulnerability in the Visakhapatnam district of India (Rao et al. 2013). Additionally, the CA was coupled with PCA for the modification of DRASTIC method in the Qazvin aquifer, in northern Iran (Javadi et al. 2017).

Thirumalaivasan et al. (2003) made a first attempt to apply AHP for the modification of DRASTIC method, which was then employed for vulnerability assessment in Tamil Nadu, India. The AHP is a structured multi-criteria analysis technique used for analyzing complex problems, and thus, the AHP is widely used for the calibration of parameters' weights in vulnerability assessment methods. In vulnerability assessment study of the Eğirdir Lake basin of Turkey, DRASTIC method was modified using the AHP technique (Sener and Davraz 2013). Furthermore, overlay and index-based techniques were coupled with the AHP in a GIS platform for estimating groundwater vulnerability in northern India (Gangadharan et al. 2016). Decision support systems (DSSs) constitute a valuable and flexible tool for groundwater resource management, and it has been also integrated with groundwater vulnerability assessment methods. For instance, the DSSs have been coupled with vulnerability maps in intensively irrigated areas of Italy and Greece providing an integrated tool for sustainable management of groundwater and optimal use of fertilizers (Voudouris et al. 2010). Gemitzi et al. (2006) developed a groundwater vulnerability index based on decision-making techniques, such as fuzzy logic and GIS. Stumpp et al. (2016) established an intrinsic vulnerability index in a decision tree form, which leads the user through the stages of vulnerability assessment. Kazakis et al. (2018a) modified the GALDIT method using fuzzy sets in order to estimate groundwater vulnerability of coastal aquifer to seawater intrusion.

The weights of evidence (WoE) method have also been used for the modification of index-based methods of groundwater vulnerability assessment. This method can evaluate the importance of each single factor class, and thus, allowing the range of values that influences the nitrate concentration to be determined. Abbasi et al. (2013) modified the DRASTIC method using WoE in the Charmahal-Bakhtyari Province in southwest Iran. In Po Plain area of Northern Italy, Sorichetta et al. (2012) used positive and negative WoE in order to assess groundwater vulnerability to nitrate. Genetic algorithm has also been used for the site selection of groundwater production wells considering groundwater vulnerability to pollution. Elçi and Ayvaz (2014) applied this approach in Tahtalı watershed in İzmir, Turkey. It is revealed from the

literature that genetic algorithm has not been widely used in studies involving assessment of the groundwater vulnerability, and therefore, this promising AI technique has the vast scope in future studies. It is worth mentioning that the applications of the aforementioned hybrid methods performed in a GIS platform highlight importance of the statistical and AI methods in development and application of the new hybrid methods. Salient hybrid methods developed by modifying the original groundwater vulnerability assessment methods are summarized in Table 2.

GIS-based groundwater quality index

The previous sections reviewed the application of time series modeling, multivariate statistical/geostatistical and artificial intelligence techniques in hydrogeochemistry, as for groundwater quality evaluation and vulnerability assessment. Combined with more conventional methods (e.g., Piper and Durov diagrams, Wilcox chart, descriptive statistics), these techniques are powerful to get a better knowledge of the geochemical processes associated to groundwater chemical evolution, both in space and time. On the other hand, a challenge remains to properly communicate the relevant geochemical knowledge to groundwater managers in a way to integrate groundwater quality issues within groundwater sustainability action framework. To do so, there was a need to develop indices that could be applied for groundwater quality assessment. Combined to GIS, these water quality indices can be integrated into other spatial data related to natural resources and human geography, thus contributing to proper development and management of the groundwater resources.

Water quality index (WQI)

In their book dealing on water quality indices (WQIs), Abbasi and Abbasi (2012) pointed out that expressing water quality brings numerous challenges. In fact, the quality of water can be defined for different uses (e.g., drinking, agricultural irrigation, livestock, and industrial), may vary in time and space, and can be categorized by a number of parameters (chemical, physical, microbiological, and radiological), with some parameters being more problematic than others, regarding health issues.

To face the complexity to describe water quality and to provide water resources managers with representation of water quality that allows comparisons between samples, regulatory agencies of different countries and international agencies have developed different types of WQIs, including the US Geological Survey (Stoner 1978), CCME-WQI (Canadian Council of Ministers of the Environment 2001), and Global Drinking (GD)-WQI (United Nations

Table 2 Methods used for groundwater vulnerability assessment and their modifications in literature

No.	Vulnerability assessment method	Type	References
1	DRASTIC	Index	Aller et al. (1985)
2	SINTACS	Index	Civita and De Maio (2004)
3	AVI	Index	Van Stempvoort et al. (1992)
4	GOD	Index	Foster (1987)
5	SEEPAGE	Index	Moore (1988)
6	Kansas leachability index	Index	Kissel et al. (1982)
7	California hotspots	Index	Cohen et al. (1986)
8	Washington map overlay vulnerability	Index	Sacha et al. (1987)
9	Iowa ground water vulnerability	Index	Hoyer and Hallberg (1991)
10	EPA/UIC	Index	Pettyjohn et al. (1991)
11	EPIK	Index	Doerfliger et al. (1999)
12	COP	Index	Zwahlen (2004)
13	PaPRIKa	Index	Kavouri et al. (2011)
14	RISKE	Index	Pételet-Giraud et al. (2000)
15	PI	Index	Goldscheider et al. (2000)
16	DRASTIC-FM	Index	Denny et al. (2007)
17	GALDIT	Index	Chachadi and Lobo-Ferreira (2001)
18	PESTANS	Simulation	Enfield et al. (1982)
19	MOUSE	Simulation	Steenhuis et al. (1987)
20	LEACHM	Simulation	Wagenet and Hutson (1987)
21	RUSTIC	Simulation	Dean et al. (1989)
22	Discriminant analysis	Statistical	Teso et al. (1988)
23	Regression analysis	Statistical	Chen and Druliner (1988)
24	Groundwater vulnerability/probability map	Hybrid	Rupert (2001)
25	DRASTIC fuzzy modification	Hybrid	Dixon (2005b)
26	DRATI	Hybrid	Panagopoulos et al. (2006)
27	DATIL	Hybrid	Antonakos and Lambrakis (2007)
28	DRASTIC fuzzy modification	Hybrid	Mohammadi et al. (2009)
29	DRASTIC fuzzy modification	Hybrid	Fijani et al. (2013)
30	DRASTIC-PAN	Hybrid	Kazakis and Voudouris (2015)
31	P3	Hybrid	Sullivan and Gao (2017)
32	DRASTIC SICM modification	Hybrid	Nadiri et al. (2017)
33	SINTACS-SV	Hybrid	Busico et al. (2017)
34	GALDIT-F	Hybrid	Kazakis et al. (2018a)
35	PaPRIKaRa	Hybrid	Kazakis et al. (2018b)

Environment Programme 2007). Lumb et al. (2011) published an extensive review of the evolution of the WQI concept, including the CCME-WQI and the GD-WQI. Later on, Sutadian et al. (2016) presented a review of 30 WQIs, developed for evaluating river water quality, based on selection of parameters, generation of sub-indices, generation of parameter weights, and aggregation procedure to compute index. Abbasi and Abbasi (2012) defined the objective of WQI as translating the concentrations of the measured parameters (variables) of a water sample into a single value (the index value). By doing so, the index value of each sample can be used to compare the water quality between samples (observations). When computing WQI with GIS, numerous applications leading to proper

and sustainable management of water resources can be implemented. Abbasi and Abbasi (2012) provide detailed information on WQI development and common generation steps, starting with the Horton's WQI (Horton 1965). Figure 4 illustrates the basic steps generally followed to develop WQI (Abbasi and Abbasi 2012). The selection of parameters (Step 1) from the water quality dataset and the attribution of a weight to each of the parameters (Step 3) are leading subjectivity to the technique. The parameters' transformation (Step 2) is needed to bring parameters of different units or ranges to a single scale, producing sub-indices. It is during this step that one can index the parameters numerically to water quality guideline, such as the World Health Organization standards (WHO 2017).

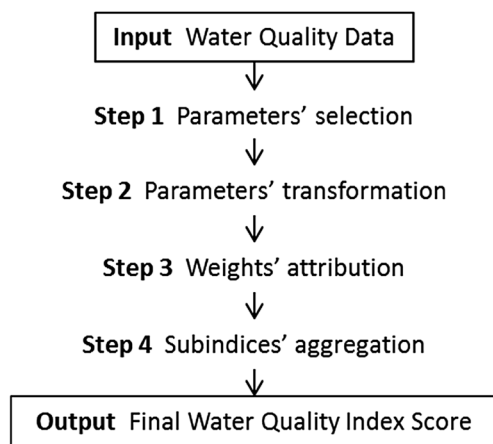


Fig. 4 Flowchart illustrating basic steps generally involved in water quality index (WQI) determination as presented by Abbasi and Abbasi (2012)

Finally, the aggregation of the sub-indices (Step 4) allows the determination of the final index score of the WQI.

Several indices were developed for assessment of surface water quality (e.g., Prati et al. 1971; Smith 1990; Dojlido et al. 1994; Nasiri et al. 2007; Thi Minh; Hanh et al. 2011; Şener et al. 2017). The following sub-section emphasizes on the application of WQI in the field of hydrogeology, where specific indices, including the groundwater quality index (GWQI; e.g., Machiwal et al. 2011), contamination index (C_d ; e.g., Backman et al. 1998), metal pollution index (MPI; e.g., Giri et al. 2010) and index of aquifer water quality (IAWQ; e.g., Melloul and Collin 1998), were developed to define the quality of groundwater. With advancement of the computing facilities, WQIs are now integrated with GIS to provide quantitative groundwater quality maps for different geographical regions and scales (e.g., Machiwal et al. 2011; Ketata et al. 2012; Sadat-Noori et al. 2014).

Groundwater quality index (GWQI) and GIS-based GWQI mapping

For the purpose of this review, GWQI is used as the general term to describe indices developed to address groundwater quality, predominantly based on physicochemical parameters (e.g., GWQI, C_d , MPI, and IAWQ). Research studies on GWQI have been reported on groundwater geochemical data from many countries, as shown in Table 3. Such studies have increased since the pioneering work of Backman et al. (1998) and Melloul and Collin (1998), with an important number of publications from semi-arid and arid regions of the world, including parts of India, where several states are facing severe water scarcity (Machiwal et al. 2011).

With the objective to provide a general view of the degree of groundwater contamination of a region, Backman et al.

(1998) tested the applicability of mapping groundwater contamination index (C_d) in Finland and Slovakia. There C_d takes into account the number of parameters exceeding the guideline values, as well as the concentrations exceeding these limit values. As shown by Backman et al. (1998), the groundwater contamination degree calculated for each sample can then be represented on maps for aesthetic and health-risk parameters distinguishing between groundwater contamination of geogenic origin and anthropogenic sources. By drawing parallels to DRASTIC model (Aller et al. 1985), Melloul and Collin (1998) developed the IAWQ, a GWQI that allowed delineating areas where land uses are already affecting groundwater quality. Stigter et al. (2006) used multivariate analysis to develop a GWQI and a groundwater composition index (GWCI) as monitoring tools for groundwater contamination from agricultural activities and to serve as communication tool as part of agro-environmental policies in Portugal.

In the last decade, several studies integrated the GWQI concept to GIS to support efficient strategies to assess groundwater quality, as well as to properly manage and monitor aquifers and groundwater resources (Table 3). Babiker et al. (2007) proposed a GIS-based GWQI with the objective to summarize available water quality data into easily understood maps. They used GIS to implement the proposed GWQI and to test the sensitivity of the model. In another GIS-based GWQI spatio-temporal study, Machiwal et al. (2011) utilized, following GWQI map, an optimum index factor (OIF) to generate a potential GWQI (P-GWQI) map in western India. They summarized the entire process to develop GWQI and P-GWQI maps within a flowchart. Following Babiker et al. (2007), Machiwal et al. (2011) also performed a map removal sensitivity analysis to identify the most influential water quality parameters, and so, the parameters that should be monitored with higher accuracy. Khan et al. (2011) used GIS-based GWQI to assess the impact of land use changes on the groundwater quality from a rapidly urbanizing region of South India. Sadat-Noori et al. (2014) combined the use of GIS and GWQI to assess groundwater quality of the Saveh-Nobaran aquifer in Iran.

GWQI has also been used in various applications in groundwater quality and hydrogeochemical studies. As an example, Ramos Leal et al. (2004) used a GWQI and the C_d of Backman et al. (1998), combined to aquifer vulnerability evaluation, to support the design of a water quality monitoring network for Mexico. Nobre et al. (2007) calculated a GWQI, based on the IAWQ of Melloul and Collin (1998), in conjunction with vulnerability, contamination and well capture indices, to develop GIS-based groundwater vulnerability and risk mapping. The approach of Nobre et al. (2007), where several elements are integrated within GIS environment, was successful to assess groundwater pollution risks and to identify areas to be prioritized for groundwater monitoring and landuse restriction. The

Table 3 Salient studies that developed and applied various types of groundwater quality indexes (GWQIs)

Country	Name of index ^a	Parameters ^b	References ^c
Brazil	GWQI	Cl, NO ₃	Nobre et al. (2007)*
	IQNAS WQI	TDS, pH, hardness, Cl, F, NO ₃	Leite et al. (2018)
China	Entropy-weighted fuzzy WQI	TDS, Ca, Mg, Na, K, Cl, SO ₄ , HCO ₃ , F, NO ₃ -N, NO ₂ -N, NH ₄ -N, Al, As, Cr, Cu, Hg, Mn, Zn	Li et al. (2014)
Croatia	GWQI/SWQI (WQI)	Temp, DO, BOD, mineralization, corrosion coefficient, total N, protein N, total P, TC	Stambuk-Giljanovic (1999)
Egypt	GWQI (WQI)	TDS, BOD, NO ₃ , Cl, PO ₄ , Cd, Cr, Ni, Pb	Soltan (1999)
	WQI	EC, TDS, pH, Na, K, hardness, SAR	El-Shahat et al. (2017)
Finland	GWQI (C _d)	pH, Na, Cl, SO ₄ , F, NO ₃ , UO ₂ , Ag, Al, As, B, Ba, Cd, Cr, Cu, Fe, Mn, Ni, Pb, Rn, Se, Zn, KMnO ₄ consumption	Backman et al. (1998)
Ghana	GWQI (WQI)	EC, Ca, Mg, Na, Cl, F, NO ₃	Banoeng-Yakubo et al. (2009)
	WQI	EC, TDS, pH, Ca, Mg, Cl, SO ₄ , HCO ₃ , NO ₃ , PO ₄	Boateng et al. (2016)
India	GWQI (WQI)	TDS, pH, hardness, Ca, Mg, Cl, SO ₄ , HCO ₃ , F, NO ₃ , Fe, Mn	Ramakrishnaiah et al. (2009)
	MPI	Cu, Fe, Mn, Ni, Pb, Zn	Giri et al. (2010)
	GWQI	EC, pH, hardness, Ca, Mg, Na, Cl, SO ₄ , alkalinity, F, NO ₃ , NO ₂ , Cd, Cr, Cu, Fe, Mn, Ni, Pb, Zn, TC, Salmonella	Ramesh et al. (2010)
	GWQI	TDS, Ca, Mg, Na, K, Cl, SO ₄ , HCO ₃ , F, NO ₃ , PO ₄ , Si	Vasanthavigar et al. (2010)
	GWQI (GQI)	TDS, Ca, Mg, Na, Cl, SO ₄	Khan et al. (2011)*
	GWQI	EC, TDS, pH, hardness, Ca, Mg, Na, Cl, SO ₄ , HCO ₃ , NO ₃	Machiwal et al. (2011)*
	GWQI (GQI)	TDS, pH, Ca, Mg, Na, K, Cl, SO ₄ , HCO ₃ , NO ₃	Sethy et al. (2017)*
Iran	GWQI	TDS, pH, Ca, Mg, Na, K, Cl, SO ₄	Saeedi et al. (2010)
	GWQI	TDS, pH, Ca, Mg, Na, K, Cl, SO ₄ , HCO ₃	Sadat-Noori et al. (2014)*
	GWQI/FGWQI	TDS, Ca, Mg, Na, Cl, SO ₄ , NO ₃	Vadiati et al. (2016)*
	EWQI	EC, Ca, Mg, Na, K, Cl, SO ₄ , HCO ₃ , F, NO ₃ , Al, As, Fe, Mn, Pb	Gorgij et al. (2017)
	PWQI	EC, TDS, pH, hardness, Ca, Mg, Na, K, Cl, SO ₄ , HCO ₃	Jamshidzadeh and Barzi (2018)
Israel	GWQI (IAWQ)	Cl, NO ₃	Melloul and Collin (1998)
Japan	GWQI (GQI)	TDS, Ca, Mg, Na, Cl, SO ₄ , NO ₃	Babiker et al. (2007)*
Lebanon	GWQI	TDS, Ca, Mg, Na, Cl, SO ₄ , HCO ₃ , F, NO ₃ , NO ₂ , FC, TC	El-Fadel et al. (2014)*
Mexico	GWCI (C _d)	Temp, EC, pH, major ions	Ramos Leal et al. (2004)
Portugal	GWQI/GWCI	Ca, Cl, SO ₄ , NO ₃	Stigter et al. (2006)
Slovakia	GWCI (C _d)	TDS, Cl, SO ₄ , F, NO ₃ , NH ₄ , Al, As, Ba, Cd, Cr, Cu, Fe, Hg, Mn, Pb, Sb, Se, Zn	Backman et al. (1998)
Syria	HPI	Cd, Cu, Pb, Zn	Zakhem and Hafez (2015)
Tunisia	GWQI	EC, TDS, pH, Ca, Mg, Na, K, Cl, SO ₄ , NO ₃	Ketata et al. (2012)*

^aGWQI/GQI groundwater quality index, IQNAS WQI groundwater natural quality index, SWQI surface water quality index, WQI water quality index, MPI metal pollution index, Cd contamination index, FGWQI fuzzy ground water quality index, EWQI entropy-weighted water quality index, PWQI potability water quality index, IAWQ index of aquifer water quality, GWCI groundwater composition index, HPI heavy metal pollution index

^bTDS total dissolved solids, temp temperature, DO dissolved oxygen, BOD biochemical oxygen demand, EC electrical conductivity, FC fecal coliform, TC total coliform, SAR sodium adsorption ratio

^cStudies with "*" have strong GIS-based GWQI applications

concept of GWQI is still in evolution, as shown by the work of Li et al. (2014) and Vadiati et al. (2016). With the objective to minimize uncertainties associated with traditional WQI calculations, Vadiati et al. (2016) investigated the potential of a hybrid fuzzy-based GWQI (FGWQI) to assess groundwater quality in the Sarab Plain of Iran. They found that the hybrid FGWQI produces significantly more accurate assessments of groundwater quality than traditional WQI. Even though the application of FGWQI is promising, Vadiati et al. (2016)

concluded that more research is needed to test the approach and compare it to deterministic WQI techniques in different contexts.

Limitations and research gaps

Whether time series analysis in hydrogeochemistry has been comprehensively applied?

This review highlighted that studies are generally lacking where multiple characteristics, i.e., normality, homogeneity, stationarity, trends, persistence, periodicity and stochasticity of a hydrogeochemical time series are characterized for the same time series. Every time series characteristics have their own importance, however, in literature studies dealing with hydrogeochemical data series, the major emphasis was only on testing normality and presence/absence of trends.

Has time series modeling been adequately integrated with GIS?

Literature reviewed in this study clearly pointed out that the studies dealing with time series of hydrogeochemical variables could not adequately integrate time series modeling with GIS technology. This is mainly due to unavailability of the essential time series analysis modules in GIS software required for analyzing time series characteristics of hydrogeochemical variables in a spatially distributed manner.

Whether all multivariate statistical techniques may adequately be coupled with GIS?

It is evident from this review that FA, PCA, and CA techniques have been extensively used to analyze the multivariate hydrogeochemical datasets. However, other MVSA techniques such as DA and CCA could not receive much attention of the researchers. In addition, it is learnt that the options for analyzing hydrogeochemical variables by applying GIS-coupled PCA technique are available in many GIS software. However, such procedures for coupling of other multivariate techniques with GIS are generally not available.

Although multivariate statistical techniques provide a powerful means for the analysis of hydrogeochemical data series (allowing simultaneous evaluation of all physico-chemical variables), due to their supervised nature, there are no unified methodologies on how to conduct such analyses. For instance, there are certain decisions to be made in certain steps of these analyses (e.g., variables to be included or excluded, use of raw or transformed/standardized variables, selection of methods, algorithms, cut-off levels, and criteria for evaluation), which may introduce a

subjective bias in the process, with the potential of greatly affecting outcomes. Even using the same dataset, different results can be obtained as a result of options chosen to conduct the analyses and during interpretation stage.

Have geostatistical-modeling techniques been advanced in mapping groundwater quality variations?

In groundwater quality mapping studies, it has been a customary practice to apply ordinary kriging technique. There have been fewer efforts in exploring applicability and efficacy of other kriging techniques, e.g., cokriging, indicator kriging, empirical Bayesian kriging, etc. Only a few studies comparatively evaluated multiple kriging techniques for getting the best results. Validation of the mapped variable is an essential step in the procedure of applying the geostatistical-modeling technique for estimating spatial distributions. However, uncertainty in spatial estimations of the groundwater quality variables remains unaddressed in few studies due to non-computation of error variance.

Is quantitative assessment of groundwater vulnerability more comprehensive than qualitative?

The modern and advanced approaches of groundwater vulnerability assessment are based on both using hybrid methods involving a coupling of qualitative–quantitative methods and integration of statistical and artificial intelligence techniques with index-based qualitative methods. However, these approaches demand a large number of field data, which are difficult to be collected/measured. Additionally, the trend to establish a method for all hydrogeological regimes might neglect the specific conditions of each aquifer. Although the qualitative approaches are characterized by subjectivity and strong dependence on the researchers, they might be more flexible, adaptable and cost-effective. It is clear that vulnerability assessment using intelligence techniques should be carefully applied and evaluated. In the future, a deeper discussion and comparison between quantitative and qualitative approaches will be needed to determine the most suitable approach.

Does GIS-based groundwater quality index provide consistent evaluation?

It is apparent from the results of past GWQI studies that index score calculation remains highly dependent on the set of parameters selected, as well as on the weights assigned to each of the parameters, making this technique subjective. Of the several developed GWQIs, hardly any GWQI have the ability to be consistent and comparable if applied over

different areas. Thus, a universally accepted and consistent GIS-based GWQI is currently not available in literature.

Are different water quality indices steady and comparable over spatial and temporal scales?

As revealed from the review of literature, most of the WQIs are dependent upon the water quality guidelines that vary among jurisdictions/institutions, and thus, lead to difficulties in interpreting water quality when comparing WQI maps at regional scale, sometimes at country-level. Furthermore, there are chances that the water quality guideline for some parameters may change in time, and in such a case, the WQI map needs to be updated prior to interpretation or use.

Perspectives on future research needs

This review clearly evidenced that application of statistical techniques such as time series modeling, multivariate statistical/geostatistical techniques, artificial intelligence techniques and water quality indices are gaining popularity in the field of hydrogeochemistry. Integration of statistical techniques with GIS enhanced capabilities for precisely interpreting the hydrogeologic processes occurring within the aquifer systems. Still, some research gaps and limitations in integration of the advanced statistical techniques within the GIS platform are experienced by the researchers as explained in the previous section, which may be a challenge for future studies undertaking water quality evaluation of the hydrogeologic systems. Few of the major needs for future research are pointed out below.

- In hydrogeochemistry, first two steps of time series modeling, i.e., detection and analysis, have been widely adopted. However, the future research will need to focus on the last two steps, i.e., synthesis and verification including stochastic time series modeling of groundwater quality variables.
- Time series characteristics other than normality and presence of trend such as homogeneity, stationarity, periodicity and persistence need to be considered equally important in hydrogeochemistry, and wide applications of statistical methods for their detection are required.
- Due to inherent complexities of and spatial continuities observed in the chemical and physical properties of water chemistry data, sometimes multivariate statistics may not be able to produce the expected results. This is mostly because, most of the multivariate statistical methods are based on binary logic (i.e., Aristotelian logic), which imposes sharp boundaries. According to this logic, a water sample can only be a member of a certain group and no overlapping groups are allowed, e.g., in cluster analysis. However, methods using a multi-valued logic (e.g., fuzzy *c*-means clustering) can be used to overcome such limitations, where partial memberships can be evaluated (a water sample can be partial member of other groups) (Güler and Thyne 2004b; Güler et al. 2012).
- In future studies, uncertainty associated with the spatial estimations of groundwater quality variables predicted by geostatistical-modeling techniques will have to be properly addressed by validating the mapped variables using the cross-validation criteria. Also, accuracy of the customary ordinary kriging and few advanced techniques such as empirical Bayesian kriging will need to be comparatively evaluated in order to find the best interpolation technique under a set of given hydrogeologic conditions.
- For groundwater vulnerability assessment studies, a need is felt to develop some sort of protocol for monitoring of the groundwater quality in order to have comparative appraisal of vulnerability degree of the aquifer over different parts of the world.
- In studies dealing with groundwater quality index (GWQI) for groundwater quality assessments, a robust methodology will have to be developed to reduce subjectivity from the processes of parameter selection and weight attribution.
- Development of a universal GWQI for assessing groundwater quality for different purposes would really be a challenging task for prospective researchers. Attempts should be made to develop a framework for generating a unique as well as versatile index that allows comparisons in groundwater quality among different spatial scales ranging from local to regional (Lumb et al. 2011).
- It is emphasized to investigate the applicability of hybrid GWQI involving coupling of artificial intelligence techniques such as fuzzy or neuro-fuzzy technique with index-based weights, along with integration of various additional parameters such as physicochemical, organic matter, microbiological, major anions/cations and heavy metals (Vadiati et al. 2016).
- One of the major future research needs will be adequate integration of all kind of statistical methods in GIS platform. Procedures for different time series modeling tests, multivariate statistical and artificial intelligence techniques and water quality indices need to be adequately incorporated in GIS software.
- Finally, it is realized that a deep sense of cooperation, sharing of experience and exchange of ideas among the hydrogeologists working in different regions of the world having a diverse setting of economic, social and political ethics would be needed to ensure thorough investigations and reliability of outcomes.

Concluding remarks

Application of modern techniques such as time series modeling, multivariate statistical/geostatistical and artificial intelligence techniques to characterize groundwater quality for efficient management and protection of groundwater resources has been attracting the researchers increasingly over the past five decades. After advent of Geographic Information System (GIS) in 1990s, the advanced statistical and artificial intelligence techniques with their GIS-based integration have emerged as more powerful tools than the traditional methods for a better evaluation of groundwater quality. It is revealed from the literature that time series modeling has not been extensively utilized in hydrogeochemistry studies, and there exists a huge scope for its comprehensive applications in future. Mainly, presence of normality and trend are examined in groundwater quality time series, and other important time series properties such as homogeneity, stationarity, periodicity, persistence and stochasticity are generally ignored on the temporal scale. Also, adequate GIS integration of time series modeling techniques are generally lacking. It is apparently depicted that the past studies have mostly adopted Factor Analysis, Principal Component Analysis (PCA) and Cluster Analysis (CA) for evaluating the groundwater quality. However, other multivariate statistical analysis techniques could not be widely employed. Recently, factor loadings/scores of the PCA have been integrated with GIS-coupled geostatistical modeling, although such studies are quite rare in literature. A need is felt to find the accuracy of the advanced geostatistical-modeling techniques, e.g., Empirical Bayesian Kriging and geostatistical-simulation, in spatial mapping of the groundwater quality variables, and their spatial comparisons with the traditional techniques, e.g., Ordinary Kriging. It is evident from literature that the concept of groundwater vulnerability has been developed widely over the past five decades since its inception in 1968, and currently a very large number of artificial intelligence techniques have seen their GIS-integrated applications for protection of the groundwater resources worldwide. In recent times, a hybrid approach amalgamating index-based rating methods with statistical methods is gaining wide attention of the researchers across the globe for groundwater vulnerability assessments. Furthermore, review of the past studies clearly highlighted that studies dealing with water quality assessment based on certain indices are relatively less for the groundwater in comparison to those for the surface water. Over the last one decade, there has been an increase in the number of studies, either developing a groundwater quality index (GWQI) or applying an existing GWQI for groundwater quality appraisals.

This review emphasizes the importance of salient time series characteristics in GIS-based hydrogeochemistry studies. In future groundwater quality studies, time series modeling should be employed in a comprehensive manner by including synthesis and verification steps. Similarly, potential of the multivariate statistical techniques, other than PCA and CA, need to be explored for groundwater quality evaluation and protection. Also, it will be imperative to minimize uncertainty of the aquifer vulnerability assessments by developing the hybrid methods with a proper balance of qualitative and quantitative methods. In addition, a robust and global GWQI would have to be generated in order to have consistent and steady water quality evaluations of the hydrogeologic systems that can be comparable over different spatial scales across the globe. Finally, one of the major future challenges would be to develop a variety of modules for implementing advanced statistical and artificial intelligence methods in GIS platform to enable advanced analyses of these techniques in a spatial manner.

Acknowledgements The authors are grateful to Dr. Olaf Kolditz (Editor-in-Chief) and two anonymous reviewers for their useful suggestions and comments, which helped improve earlier version of this article.

References

- Abbasi T, Abbasi SA (2012) Water quality indices. Elsevier, Oxford, UK, p 384
- Abbasi S, Mohammadi K, Kholghi MK, Howard K (2013) Aquifer vulnerability assessments using DRASTIC, weights of evidence and the analytic element method. *Hydrol Sci J* 58(1):186–197
- Adeloye AJ, Montaseri M (2002) Preliminary streamflow data analyses prior to water resources planning study. *Hydrol Sci J* 47(5):679–692
- Adhikary PP, Dash CJ, Chandrasekharan H, Rajput TBS, Dubey SK (2012) Evaluation of groundwater quality for irrigation and drinking using GIS and geostatistics in a peri-urban area of Delhi, India. *Arab J Geosci* 5:1423–1434
- Ağca N, Karanlık S, Ödemiş B (2014) Assessment of ammonium, nitrate, phosphate, and heavy metal pollution in groundwater from Amik Plain, southern Turkey. *Environ Monit Assess* 186(9):5921–5934
- Aguilar JB, Orban P, Dassargues A, Brouyère S (2007) Identification of groundwater quality trends in a chalk aquifer threatened by intensive agriculture in Belgium. *Hydrogeol J* 15:1615–1627
- Ahmed S (2006) Application of Geostatistics in Hydrosciences. In: Thangarajan M (ed) Groundwater resource evaluation, augmentation, contamination, restoration, modeling and management. Capital Publishing Company, New Delhi, pp 78–111
- Aller L, Bennett T, Lehr JH, Petty RJ (1985) DRASTIC: a standardized system for evaluating ground water potential using hydrogeological settings. In: EPA/600/285/018, Environmental Research Laboratory. US Environmental Protection Agency, Ada
- Alley WM (1993) Regional ground-water quality. Wiley, New York, p 634
- Alther GA (1979) A simplified statistical sequence applied to routine water quality analysis—a case history. *Ground Water* 17(6):556–561

- Antonakos AK, Lambrakis NL (2007) Development and testing of three hybrid methods for assessment of aquifer vulnerability to nitrates, based on the DRASTIC model, an example from NE Korinthia, Greece. *J Hydrol* 333(2–4):288–304
- APHA-AWWA-WEF (2017) Standard methods for the examination of water and wastewater, 23rd edn. American Public Health Association (APHA)-American Water Works Association (AWWA)-Water Environment Federation (WEF), Washington, DC
- Arslan H (2017) Determination of temporal and spatial variability of groundwater irrigation quality using geostatistical techniques on the coastal aquifer of Çarşamba Plain, Turkey, from 1990 to 2012. *Environ Earth Sci* 76:38. <https://doi.org/10.1007/s12665-016-6375-x>
- ASCE Task Committee (1990a) Review of geostatistics in geohydrology. I: Basic concepts. *Journal of Hydraul Eng ASCE* 116(5):612. [https://doi.org/10.1061/\(ASCE\)0733-9429\(1990\)116:5\(612\)](https://doi.org/10.1061/(ASCE)0733-9429(1990)116:5(612))
- ASCE Task Committee (1990b) Review of geostatistics in geohydrology. II: applications. *J Hydraul Eng ASCE* 116(5):612. [https://doi.org/10.1061/\(ASCE\)0733-9429\(1990\)116:5\(633\)](https://doi.org/10.1061/(ASCE)0733-9429(1990)116:5(633))
- ASCE Task Committee (2000) Artificial neural networks in hydrology—I: preliminary concepts. *J Hydrol Eng ASCE* 5(2):115–123
- Ashley RP, Lloyd JW (1978) An example of the use of factor analysis and cluster analysis in groundwater chemistry interpretation. *J Hydrol* 39:355–364
- Assaf H, Saadeh M (2009) Geostatistical assessment of groundwater nitrate contamination with reflection on DRASTIC vulnerability assessment: the case of the Upper Litani Basin, Lebanon. *Water Resour Manag* 23:775–796
- Ayuso SV, Acebes P, López-Archilla AI, Montes C, Guerrero MC (2009) Environmental factors controlling the spatiotemporal distribution of microbial communities in a coastal, sandy aquifer system (Doñana, southwest Spain). *Hydrogeol J* 17:767–780
- Babiker IS, Mohamed MMA, Hiyama T (2007) Assessing groundwater quality using GIS. *Water Resour Manag* 21:699–715
- Backman B, Bodiš D, Lahermo P, Rapant S, Tarvainen T (1998) Application of a groundwater contamination index in Finland and Slovakia. *Environ Geol* 36(1–2):55–64
- Banoeng-Yakubo B, Yidana SM, Emmanuel N, Akabzaa T, Asiedu D (2009) Analysis of groundwater quality using water quality index and conventional graphical methods: the Volta region, Ghana. *Environ Earth Sci* 59(4):867–879
- Barca E, Passarella G (2008) Spatial evaluation of the risk of groundwater quality degradation. A comparison between disjunctive kriging and geostatistical simulation. *Environ Monit Assess* 137:261–273
- Bárdossy A (2011) Interpolation of groundwater quality parameters with some values below the detection limit. *Hydrol Earth Syst Sci* 15:2763–2775
- Bárdossy A, Kundzewicz ZW (1990) Geostatistical methods for detection of outliers in groundwater quality spatial fields. *J Hydrol* 115:343–359
- Berzas JJ, Garcia LF, Rodriguez RC, Martinalvarez PJ (2000) Evolution of the water quality of a managed natural wetland: Tablas de Daimiel National Park (Spain). *Water Res* 34(12):3161–3170
- Bethea RM, Rhinehart RR (1991) Applied engineering statistics. Marcel Dekker, Inc., New York
- Bhuiyan MAH, Bodrud-Doza M, Islam ARMT, Rakib MA, Rahman MS, Ramanathan AL (2016) Assessment of groundwater quality of Lakshimpur district of Bangladesh using water quality indices, geostatistical methods, and multivariate analysis. *Environ Earth Sci* 75:1020. <https://doi.org/10.1007/s12665-016-5823-y>
- Bjerg PL, Christensen TH (1992) Spatial and temporal small-scale variation in groundwater quality of a shallow sandy aquifer. *J Hydrol* 131:133–149
- Boateng TK, Opoku F, Acquah SO, Akoto O (2016) Groundwater quality assessment using statistical approach and water quality index in Ejisu-Juaben Municipality, Ghana. *Environ Earth Sci* 75:489. <https://doi.org/10.1007/s12665-015-5105-0>
- Bodrud-Doza M, Islam ARMT, Ahmed F, Das S, Saha N, Rahman MS (2016) Characterization of groundwater quality using water evaluation indices, multivariate statistics and geostatistics in central Bangladesh. *Water Sci* 30:19–40
- Bondu R, Cloutier V, Rosa E, Benzaazoua M (2016) A review and evaluation of the impacts of climate change on geogenic arsenic in groundwater from fractured bedrock aquifers. *Water Air Soil Pollut* 227:296. <https://doi.org/10.1007/s11270-016-2936-6>
- Bondu R, Cloutier V, Rosa E, Benzaazoua M (2017) Mobility and speciation of geogenic arsenic in bedrock groundwater from the Canadian Shield in western Quebec, Canada. *Sci Total Environ* 574:509–519
- Box GEP, Cox DR (1964) An analysis of transformations. *J R Stat Soc Ser B* 26(2):211–252
- Box GEP, Hunter WG, Hunter JS (1978) Statistics for experimenters: an introduction to design, data analysis, and model building. Wiley Interscience, New York
- Boy-Roura M, Nolan BT, Menció A, Mas-Pla J (2013) Regression model for aquifer vulnerability assessment of nitrate pollution in the Osona region (NE Spain). *J Hydrol* 505:150–162
- Bras RL, Rodriguez-Iturbe I (1985) Random functions and hydrology. Addison-Wesley, Reading
- Bronson KF, Malapati A, Booker JD, Scanlon BR, Hudnall WH, Schurbert AM (2009) Residual soil nitrate in irrigated Southern High Plains cotton fields and Ogallala groundwater nitrate. *J Soil Water Conserv* 64(2):98–104
- Brown CE (1998) Applied multivariate statistics in geohydrology and related sciences, 1st edn. Springer, New York
- Buragohain M, Bhuyan B, Sarma HP (2010) Seasonal variations of lead, arsenic, cadmium and aluminium contamination of groundwater in Dhemaji district, Assam, India. *Environ Monit Assess* 170:345–351
- Burrough PA, McDonnell RA (1998) Principles of geographical information systems. Oxford University Press, Oxford, 333 pp
- Busico G, Kazakis N, Colombani N, Mastrocicco M, Voudouris K, Tedesco D (2017) A modified SINTACS method for groundwater vulnerability and pollution risk assessment in highly anthropized regions based on NO_3^- and SO_4^{2-} concentrations. *Sci Total Environ* 609:1512–1523
- Busico G, Cuoco E, Kazakis N, Colombani N, Mastrocicco M, Tedesco D, Voudouris K (2018) Multivariate statistical analysis to characterize/discriminate between anthropogenic and geogenic trace element occurrence in Campania Plain, Southern Italy. *Environ Pollut* 234:260–269
- Cairns SH, Dickson KL, Atkinson SF (1997) An examination of measuring selected water quality trophic indicators with SPOT satellite HRV data. *Photogramm Eng Remote Sens* 63(3):263–265
- Canadian Council of Ministers of the Environment (2001) Canadian water quality guidelines for the protection of aquatic life. CCME Water Quality Index 1.0, Technical Report, in Canadian Environmental Quality Guidelines, 1999, Canadian Council of Ministers of the Environment, Winnipeg, Canada
- Candela L, Olea RA, Custodio E (1988) Lognormal kriging for the assessment of reliability in groundwater quality control observation networks. *J Hydrol* 103:67–84
- Castrignanò A, Giugliarini L, Risaliti R, Martinelli N (2000) Study of spatial relationships among some soil physicochemical properties of a field in central Italy using multivariate geostatistics. *Geoderma* 97:39–60
- Chachadi AG, Lobo-Ferreira JP (2001) Seawater intrusion vulnerability mapping of aquifers using the GALDIT method. In: Proceedings of the workshop on modelling in hydrogeology, Anna University, Chennai, pp 143–156

- Chang K-T (2002) Introduction to geographic information systems. Tata McGraw-Hill Publishing Company Ltd., New Delhi, p 348
- Chaudhuri S, Ale S, Delaune P, Rajan N (2012) Spatio-temporal variability in groundwater nitrate concentration in Texas: 1960–2010. *J Environ Qual* 41:1806–1817
- Chen H, Druliner AD (1988) Agricultural chemical contamination of ground water in six areas of the high plains aquifer, Nebraska. National Water Summary 1986—hydrologic events and ground-water quality, water-supply Paper 2325. U.S. Geological Survey, Reston
- Chen L, Feng Q (2013) Geostatistical analysis of temporal and spatial variations in groundwater levels and quality in the Minqin Oasis, Northwest China. *Environ Earth Sci* 70(3):1367–1378
- Chen Y, Takara K, Cluckie ID, Smedt FHD (eds) (2004) GIS and remote sensing in hydrology, Water resources and environment. IAHS Publication No. 289. IAHS Press, Wallingford, p 422
- Chou CJ (2006) Assessing spatial, temporal, and analytical variation of groundwater chemistry in a large nuclear complex, USA. *Environ Monit Assess* 119:571–598
- Civita M, De Maio M (2004) Assessing and mapping groundwater vulnerability to contamination: the Italian “combined” approach. *Geofísica Internacional* 43(4):513–532
- Clark D (1975) Understanding canonical correlation analysis. Concepts and techniques in modern geography No. 3, geo abstracts. University of East Anglia, Norwich
- Clarke RT (1998) Stochastic processes for water scientists: development and applications. Wiley, New York
- Cloutier V, Lefebvre R, Therrien R, Savard MM (2008) Multivariate statistical analysis of geochemical data as indicative of the hydrogeochemical evolution of groundwater in a sedimentary rock aquifer system. *J Hydrol* 353:294–313
- Cohen DB, Fisher C, Reid ML (1986) Ground-water contamination by toxic substances: a California assessment. In: Garner WY, Honeycutt RC, Nigg HN (eds) Evaluation of pesticides in ground water, ACS Symposium Series 315. American Chemical Society, Washington, DC, pp. 499–529
- Collins WD (1923) Graphic representation of water analyses. *Ind Eng Chem* 15(4):394
- Cooley WW, Lohnes PR (1971) Multivariate data analysis. Wiley, New York
- Cooper RM, Istok JD (1988a) Geostatistics applied to groundwater contamination. I: methodology. *J Environ Eng ASCE* 111(2):270–286
- Cooper RM, Istok JD (1988b) Geostatistics applied to groundwater contamination. II. *Appl J Environ Eng ASCE* 111(2):287–299
- Council of Canadian Academies (2009) The sustainable management of groundwater in Canada. Report of the expert panel on groundwater. Council of Canadian Academies, Ottawa
- Cryer JD (1986) Time series analysis. PWS Publishers, Duxbury Press, Boston
- D’Agostino V, Greene EA, Passarella G, Vurro M (1998) Spatial and temporal study of nitrate concentration in groundwater by means of coregionalization. *Environ Geol* 36(3–4):285–295
- Dalton MG, Upchurch SB (1978) Interpretation of hydrochemical facies by factor analysis. *Ground Water* 16(4):228–233
- David M (1977) Geostatistical ore reserve estimation. Elsevier Scientific Publishing Company, New York, p 364
- David M, Dagbert M (1975) Lakeview revisited: variograms and correspondence analysis—new tools for the understanding of geochemical data. In: Proceedings of the 5th international geochemical exploration symposium, geochemical exploration, pp 163–181
- Davies PJ, Crosbie RS (2018) Mapping the spatial distribution of chloride deposition across Australia. *J Hydrol* 561:76–88
- Davis JC (1986) Statistics and data analysis in geology, 2nd edn. Wiley, New York
- Dawdy DR, Feth JH (1967) Applications of factor analysis in study of chemistry of ground water quality, Mojave River Valley, California. *Water Resour Res* 3(2):505–510
- Dean JD, Huyakorn PS, Donigian AS Jr, Voos KA, Schanz RW, Meeks YJ, Carsel RF (1989) Risk of unsaturated/saturated transport and transformation of chemical concentrations (RUSTIC). Volumes I and II. EPA/600/3–89/048a. United States Environmental Protection Agency, Athens
- Delhomme JP (1978) Kriging in hydrosociences. *Adv Water Resour* 1:251–266
- Denny SC, Allen DM, Journeay JM (2007) DRASTIC-Fm: a modified vulnerability mapping method for structurally controlled aquifers in the southern Gulf Islands, British Columbia, Canada. *Hydrogeol J* 15:483–493
- Deutsch WJ (1997) Groundwater geochemistry: fundamentals and applications to contamination. CRC Press LLC, Boca Raton, p 221
- Dillon R, Goldstein M (1984) multivariate analyses: methods and applications. Wiley, New York
- Dixon B (2005a) Groundwater vulnerability mapping: A GIS and fuzzy rule based integrated tool. *J Appl Geogr* 25:327–347
- Dixon B (2005b) Applicability of neuro-fuzzy techniques in predicting groundwater vulnerability: a GIS-based sensitivity analysis. *J Hydrol* 309(1–4):17–38
- Dixon B (2009) A case study using support vector machines, neural networks and logistic regression in a GIS to identify wells contaminated with NO₃-N. *Hydrogeol J* 17:1507–1520
- Doerfliger N, Jeannin PY, Zwahlen F (1999) Water vulnerability assessment in karst environments: a new method of defining protection areas using a multi-attribute approach and GIS tools (EPIK method). *Environ Geol* 39(2):165–176
- Dojlido J, Raniszewsk IJ, Woyciechowska J (1994) Water quality index—application for rivers in Vistula river basin in Poland. *Water Sci Technol* 30:57–64
- Dokou Z, Kourgiyalas NN, Karatzas GP (2015) Assessing groundwater quality in Greece based on spatial and temporal analysis. *Environ Monit Assess* 187:774. <https://doi.org/10.1007/s10661-015-4998-0>
- Dragon K (2006) Application of factor analysis to study contamination of a semi-confined aquifer (Wielkopolska Buried Valley aquifer, Poland). *J Hydrol* 331:272–279
- Drozdzov OA, Shepelevskii AA (1946) The theory of interpolation in a stochastic field of meteorological elements and its application to meteorological maps and network rationalization problems (in Russian). Trudy NIU GUGMS Series, 1(13), Russian Hydrological and Meteorological Service, Russia
- Eheart JW, Cieniawski SE, Ranjithan S (1993) Genetic-algorithm-based design of groundwater quality monitoring system. WRC Research Report No. 218, Water Resources Center, University of Illinois at Urbana-Champaign, 205 North Mathews. Avenue Urbana Illinois 61801, p 50
- Elçi A, Ayvaz MT (2014) Differential-evolution algorithm based optimization for the site selection of groundwater production wells with the consideration of the vulnerability concept. *J Hydrol* 511:736–749
- El-Fadel M, Tomaszewicz M, Adra Y, Sadek S, Najm MA (2014) GIS-based assessment for the development of a groundwater quality index towards sustainable aquifer management. *Water Resour Manag* 28:3471–3487
- El-Shahat MF, Sadek MA, Embaby AA, Salem WM, Mohamed FA (2017) Hydrochemical and multivariate analysis of groundwater quality in the northwest of Sinai, Egypt. *J Water Health*. <https://doi.org/10.2166/wh.2017.276>
- Enfield CG, Carsel RF, Cohen SZ, Phan T, Walters DM (1982) Approximating pollutant transport to ground water. *Ground Water* 20(6):711–722

- Enwright N, Hudak PF (2009) Spatial distribution of nitrate and related factors in the high plains aquifer. *Texas Environ Geol* 58:1541–1548
- Farnham M, Klaus JS, Ashok KS, Johannesson KH (2000) Deciphering groundwater flow systems in Oasis Valley, Nevada, using trace element chemistry, multivariate statistics, and geographical information system. *Math Geol* 32(8):943–968
- Farnham M, Singh AK, Stetzenbach KJ, Johannesson KH (2002) Treatment of nondetects in multivariate analysis of groundwater geochemistry data. *Chemometr Intell Lab Syst* 60:265–281
- Felmy AR, Girvin DC, Jenne EA (1983) MINTEQ: A Computer Program for Calculating Aqueous Geochemical Equilibria. EPA/600/3–84/032, Pacific Northwest Laboratory, United States Environmental Protection Agency (USEPA), Washington, DC
- Ferguson G, Gleeson T (2012) Vulnerability of coastal aquifers to groundwater use and climate change. *Nat Clim Change* 2:342–345
- Fijani E, Nadiri AA, Moghaddam AA, Tsai FT-C, Dixon B (2013) Optimization of DRASTIC method by supervised committee machine artificial intelligence to assess groundwater vulnerability for Maragheh-Bonab plain aquifer, Iran. *J Hydrol* 503:89–100
- Forina M, Armanino C, Raggio V (2002) Clustering with dendrograms on interpretation variables. *Anal Chim Acta* 454(1):13–19
- Foster SSD (1987) Fundamental concepts in aquifer vulnerability, pollution risk and protection strategy. In: Van Duijvenbooden W, and Waegeningh HG (eds), *Vulnerability of soil and groundwater to pollutants*. In: TNO committee on hydrological research, the Hague, Proc. Inf., vol 38, pp 69–86
- Frans L (2008) Trends of pesticides and nitrate in ground water of the Central Columbia Plateau, Washington, 1993–2003. *J Environ Qual* 37:273–280
- Freeze RA, Cherry JA (1979) *Groundwater*. Prentice-Hall, Inc., Englewood Cliffs
- Gan Y, Zhao K, Deng Y, Liang X, Ma T, Wang Y (2018) Groundwater flow and hydrogeochemical evolution in the Jiangnan Plain, central China. *Hydrogeol J* 26(5):1609–1623
- Gandin LS (1965) Objective analysis of meteorological fields. *Israel Program for Scientific Translations, Jerusalem*, p 242
- Gangadharan R, Nila Rekha P, Vinoth S (2016) Assessment of groundwater vulnerability mapping using AHP method in coastal watershed of shrimp farming area. *Arab J Geosci* 9: 107. <https://doi.org/10.1007/s12517-015-2230-8>
- Gemitzi A, Petalas C, Tsihrintzis VA, Pisinaras V (2006) Assessment of groundwater vulnerability to pollution: a combination of GIS, fuzzy logic and decision making techniques. *Environ Geol* 49:653–673
- Gibbs RJ (1970) Mechanisms controlling world water chemistry. *Science* 170:1088–1090
- Giggenbach WF (1988) Geothermal solute equilibria. Derivation of Na-K-Mg-Ca geoindicators. *Geochim Cosmochim Acta* 52(12):2749–2765
- Gilbert RO (1987) *Statistical methods for environmental pollution monitoring*. Van Nostrand Reinhold, New York
- Giles BD, Flocas AA (1984) Air temperature variation in Greece, Part I: persistence, trend and fluctuations. *Int J Climatol* 4:531–539
- Giri S, Singh G, Gupta SK, Jha VN, Tripathi RM (2010) An evaluation of metal contamination in surface and groundwater around a proposed uranium mining site, Jharkhand, India. *Mine Water Environ* 29(3):225–234
- Gleeson T, VanderSteen J, Sophocleous MA, Taniguchi M, Alley WM, Allen DM, Zhou Y (2010) Groundwater sustainability strategies. *Nat Geosci* 3:378–379
- Gogu RC, Dassargues A (2000) Current trends and future challenges in groundwater vulnerability assessment using overlay and index methods. *Environ Geol* 39(6):549–559
- Goldscheider N, Klute M, Sturm S, Hotzl H (2000) The PI method—a GIS-based approach to mapping groundwater vulnerability with special consideration of karst aquifers. *Z Angew Geol* 46(3):157–166
- Gong X, Richman MB (1995) On the application of cluster analysis to growing season precipitation data in North America East of the Rockies. *J Clim* 8:897–931
- Goodchild MF, Parks BO, Steyaert LT (eds) (1993) *Environmental modeling with GIS*. Oxford University Press, New York
- Goovaerts P (1999) *Geostatistics in soil science: state-of-the-art and perspectives*. *Geoderma* 89:1–45
- Goovaerts P, AvRuskin G, Meliker J, Slotnick M, Jacquez G, Nriagu J (2005) Geostatistical modeling of the spatial variability of arsenic in groundwater of southeast Michigan. *Water Resour Res* 41:W07013. <https://doi.org/10.1029/2004WR003705>
- Gorgij AD, Kisi O, Moghaddam AA, Taghipour A (2017) Groundwater quality ranking for drinking purposes, using the entropy method and the spatial autocorrelation index. *Environ Earth Sci* 76:269. <https://doi.org/10.1007/s12665-017-6589-6>
- Güler C, Thyne GD (2004a) Hydrologic and geologic factors controlling surface and groundwater chemistry in Indian Wells-Owens Valley area, southeastern California, USA. *J Hydrol* 285(1–4):177–198
- Güler C, Thyne GD (2004b) Delineation of hydrochemical facies distribution in a regional groundwater system by means of fuzzy c-means clustering. *Water Resour Res* 40(12):W12503. <https://doi.org/10.1029/2004WR003299>
- Güler C, Thyne GD, McCray JE, Turner AK (2002) Evaluation of graphical and multivariate statistical methods for classification of water chemistry data. *Hydrogeol J* 10(4):455–474
- Güler C, Kurt MA, Alpaslan M, Akbulut C (2012) Assessment of the impact of anthropogenic activities on the groundwater hydrology and chemistry in Tarsus coastal plain (Mersin, SE Turkey) using fuzzy clustering, multivariate statistics and GIS techniques. *J Hydrol* 414–415:435–451
- Güler C, Thyne GD, Tağa H, Yıldırım Ü (2017) Processes governing alkaline groundwater chemistry within a fractured rock (ophiolitic mélange) aquifer underlying a seasonally inhabited headwater area in the Aladağlar Range (Adana, Turkey). *Geofluids*. <https://doi.org/10.1155/2017/3153924> (article ID 3153924)
- Gupta A, Kamble T, Machiwal D (2017) Comparison of ordinary and Bayesian kriging techniques in depicting rainfall variability in arid and semi-arid regions of northwest India. *Environ Earth Sci* 76:512. <https://doi.org/10.1007/s12665-017-6814-3>
- Gurdak JJ, McMahon PB, Bruce BW (2012) Vulnerability of groundwater quality to human activity and climate change and variability, high plains aquifer, USA. In: Treidel H, Martin-Bordes JL, Gurdak JJ (eds) *Climate Change effects on groundwater resources—a global synthesis of findings and recommendations*. Taylor & Francis Group, London, pp 145–168
- Gurnell AM, Montgomery DR (eds) (2000) *Hydrological applications of GIS*. Wiley, Chichester, p 176
- Haan CT (1977) *Statistical methods in hydrology*. Iowa State University Press, Iowa
- Hanh TM, Sthiannopkao P, The Ba S, D. and Kim K-W (2011) Development of water quality indexes to identify pollutants in Vietnam's surface water. *J Environ Eng ASCE* 137(4):273–283
- Hardy A (1996) On the number of clusters. *Comput Stat Data Anal* 23:83–96
- Hartigan A (1975) *Clustering algorithms*. Wiley, New York
- Hassan MM, Atkins PJ (2007) Arsenic risk mapping in Bangladesh: A simulation technique of cokriging estimation from regional count data. *J Environ Sci Health Part A Toxic/Hazard Substanc Environ Engin* 42(12):1719–1728
- Helsel DR, Frans LM (2006) Regional Kendall test for trend. *Environ Sci Technol* 40(13):4066–4073

- Helstrup T, Jørgensen NO, Banoeng-Yakubo B (2007) Investigation of hydrochemical characteristics of groundwater from the Cretaceous-Eocene limestone aquifer in southern Ghana and southern Togo using hierarchical cluster analysis. *Hydrogeol J* 15:977–989
- Hem JD (1970) Study and interpretation of the chemical characteristics of natural water, 2nd edition, United States Geological Survey Water-Supply Paper 1473, Washington DC
- Hem JD (1985) Study and Interpretation of the chemical characteristics of natural water. 3rd edition, United States Geological Survey Water-Supply Paper 2254, Washington DC
- Horton RK (1965) An index number system for rating water quality. *J Water Pollut Control Fed* 37(3):300–306
- Hosseini SM, Mahjouri N (2014) Developing a fuzzy neural network-based support vector regression (FNN-SVR) for regionalizing nitrate concentration in groundwater. *Environ Monit Assess* 186:3685–3699
- Hotelling H (1936) Relations between two sets of variates. *Biometrika* 28:312–377
- Hoyer BE, Hallberg GR (1991) Ground water vulnerability regions of Iowa, special map 11. Iowa Department of Natural Resources, Iowa City
- Hudak PF (2000a) Sulfate and chloride concentrations in Texas aquifer. *Environ Int* 26:55–61
- Hudak PF (2000b) Regional trends in nitrate content of Texas groundwater. *J Hydrol* 228:37–47
- Hudak PF (2001) Water hardness and sodium trends in Texas aquifers. *Environ Monit Assess* 68:177–185
- Iskandar I, Koike K (2011) Distinguishing potential sources of arsenic released to groundwater around a fault zone containing a mine site. *Environ Earth Sci* 63:595–608
- Iskandar I, Koike K, Sendjaja P (2012) Identifying groundwater arsenic contamination mechanisms in relation to arsenic concentrations in water and host rocks. *Environ Earth Sci* 65:2015–2026
- Istok JD, Cooper RM (1988) Geostatistics applied to groundwater contamination. III: global estimates. *J Environ Engin ASCE* 111(2):915–928
- Istok JD, Rautman CA (1996) Probabilistic assessment of groundwater contamination: 2. Results of case study. *Ground Water* 34(6):1050–1064
- Istok JD, Smyth JD, Flint FL (1993) Multivariate geostatistical analysis of ground-water contamination: a case history. *Ground Water* 31(1):63–74
- Izenman AJ (2013) Modern multivariate statistical techniques: regression, classification, and manifold learning, 2nd edn. Springer, New York
- Jacobs J, Testa S (2004) Overview of chromium (VI) in the environment: background and history. In: Guertin J, Jacobs J, Avakian C (eds) Chromium (VI) handbook. CRC Press, New York. [http://www.engr.uconn.edu/~baholmen/docs/ENVE290W/National%20Chromium%20Files%20From%20Luke/Cr\(VI\)%20Handbook/L1608_C01.pdf](http://www.engr.uconn.edu/~baholmen/docs/ENVE290W/National%20Chromium%20Files%20From%20Luke/Cr(VI)%20Handbook/L1608_C01.pdf). Accessed 21 June 2017
- Jamshidzadeh Z, Barzi MT (2018) Groundwater quality assessment using the potability water quality index (PWQI): a case in the Kashan plain, Central Iran. *Environ Earth Sci* 77:59. <https://doi.org/10.1007/s12665-018-7237-5>
- Jang C-S (2013) Use of multivariate indicator kriging methods for assessing groundwater contamination extents for irrigation. *Environ Monit Assess* 185:4049–4061
- Jang C-S, Liu C-W, Lu KL, Lin CC (2007) Delimitation of arsenic-contaminated groundwater using risk-based indicator approaches around blackfoot disease hyperendemic areas of southern Taiwan. *Environ Monit Assess* 134:293–304
- Jang C-S, Chen C-F, Liang C-P, Chen J-S (2016) Combining groundwater quality analysis and a numerical flow simulation for spatially establishing utilization strategies for groundwater and surface water in the Pingtung Plain. *J Hydrol* 533:541–556
- Javadi S, Hashemy SM, Mohammadi K, Howard KWF, Neshat A (2017) Classification of aquifer vulnerability using K-means cluster analysis. *J Hydrol* 549:27–37
- Jha MK, Chowdhury A, Chowdary VM, Peiffer S (2007) Groundwater management and development by integrated remote sensing and geographic information systems: prospects and constraints. *Water Resour Manag* 21(2):427–467
- Johnson RA, Wichern DW (1992) Applied multivariate statistical analysis, 3rd edn. Prentice-Hall International, Englewood Cliffs, p 642
- Jones AL, Smart PL (2005) Spatial and temporal changes in the structure of groundwater nitrate concentration time series (1935–1999) as demonstrated by autoregressive modeling. *J Hydrol* 310:201–215
- Journel A (1974) Geostatistics for conditional simulation of orebodies. *Econ Geol* 69(5):673–687
- Journel AG, Huijbregts CJ (1978) Mining geostatistics. Academic Press, London, p 600
- Jovein EB, Hosseini SM (2017) Predicting saltwater intrusion into aquifers in vicinity of deserts using spatio-temporal kriging. *Environ Monit Assess* 189:81. <https://doi.org/10.1007/s10661-017-5795-8>
- Kagan RL (1967) Some problems relative to the interpretation of rainfall data (in Russian). *Trudy GGO* 208:64–75
- Kallio MP, Mujunen SP, Hatzimihalis G, Koutoufides P, Minkkinen P, Wilki PJ, Connor MA (1999) Multivariate data analysis of key pollutants in sewage samples: a case study. *Anal Chim Acta* 393(1–3):181–191
- Kaown D, Hyun Y, Bae G-O, Oh CW, Lee K-K (2012) Evaluation of spatio-temporal trends of groundwater quality in different land uses using Kendall test. *Geosci J* 16(1):65–75
- Karami S, Madani H, Katibeh H, Marj AF (2018) Assessment and modeling of the groundwater hydrogeochemical quality parameters via geostatistical approaches. *Appl Water Sci* 8:23. <https://doi.org/10.1007/s13201-018-0641-x>
- Karant K (1987) Ground water assessment: development and management. Tata McGraw-Hill Publishing Company Limited, New Delhi, p 720
- Kavouri K, Plagnes V, Tremoulet J, Dorfliger N, Rejiba F, Marchet P (2011) PaPRIKa: a method for estimating karst resource and source vulnerability—application to the Ouyse karst system (southwest France). *Hydrogeol J* 19:339–353
- Kazakis N, Voudouris KS (2015) Groundwater vulnerability and pollution risk assessment of porous aquifers to nitrate: modifying the DRASTIC method using quantitative parameters. *J Hydrol* 525:13–25
- Kazakis N, Kantiranis N, Voudouris KS, Mitrakas M, Kaprara E, Pavlou A (2015) Geogenic Cr oxidation on the surface of mafic minerals and the hydrogeological conditions influencing hexavalent chromium concentrations in groundwater. *Sci Total Environ* 514:224–238
- Kazakis N, Kantiranis N, Kalaitzidou K, Kaprara E, Mitrakas M, Frei R, Vargemezis G, Tsourlos P, Zouboulis A, Filippidis A (2017) Origin of hexavalent chromium in groundwater: the example of Sarigkiol Basin, Northern Greece. *Sci Total Environ* 593–594:552–566
- Kazakis N, Spiliotis M, Voudouris K, Pliakas FK, Papadopoulos B (2018a) A fuzzy multicriteria categorization of the GALDIT method to assess seawater intrusion vulnerability of coastal aquifers. *Sci Total Environ* 621:552–566
- Kazakis N, Chalikakis K, Mazzilli N, Ollivier C, Manakos A, Voudouris K (2018b) Management and research strategies of karst aquifers in Greece: Literature overview and exemplification based on hydrodynamic modelling and vulnerability assessment of a strategic karst aquifer. *Sci Total Environ* 643:592–609
- Kendall MG (1973) Time series. Charles Griffin and Co. Ltd., London

- Ketata M, Gueddari M, Bouhlila R (2012) Use of geographical information system and water quality index to assess groundwater quality in El Khairat deep aquifer (Enfidha, Central East Tunisia). *Arab J Geosci* 5:1379–1390
- Khan HH, Khan A, Ahmed S, Perrin J (2011) GIS-based impact assessment of land-use changes on groundwater quality: study from a rapidly urbanizing region of South India. *Environ Earth Sci* 63:1289–1302
- Kim TH, Chung SY, Park N, Hamm S-Y, Lee SY, Kim B-W (2012) Combined analyses of chemometrics and kriging for identifying groundwater contamination sources and origins at the Masan coastal area in Korea. *Environ Earth Sci* 67(5):1373–1388
- Kissel DE, Bidwell OW, Kientz JF (1982) Leaching classes in Kansas Soils. *Bulletin No. 641*. Kansas State University, Manhattan
- Koh E-H, Lee SH, Kaown D, Moon HS, Lee E, Lee K-K, Kang BR (2017) Impacts of land use change and groundwater management on long-term nitrate-nitrogen and chloride trends in groundwater of Jeju Island, Korea. *Environ Earth Sci*, 76: 176. <https://doi.org/10.1007/s12665-017-6466-3>
- Konikow L, Kendy L (2005) Groundwater depletion: a global problem. *Hydrogeol J* 13(1):317–320
- Kumar S, Machiwal D, Dayal D (2017) Spatial modeling of rainfall trends using satellite datasets and geographic information system. *Hydrol Sci J* 62(10):1636–1653
- Kurumbain WC, Graybill FA (1965) An introduction to statistical models in geology. McGraw-Hill, New York
- Lambrakis N, Antonakos A, Panagopoulos G (2004) The use of multi-component statistical analysis in hydrogeological environmental research. *Water Res* 38:1862–1872
- Langelier W, Ludwig H (1942) Graphical methods for indicating the mineral character of natural waters. *J Am Water Assoc* 34:335–352
- Leite NK, Stolberg J, da Cruz SP, de Tavela OA, Safanelli JL, Marchini HR, Exterkoetter R, Leite GMC, Krusche AV, Johnson MS (2018) Hydrochemistry of shallow groundwater and springs used for potable supply in Southern Brazil. *Environ Earth Sci* 77:80. <https://doi.org/10.1007/s12665-018-7254-4>
- Li P, Wu J, Qian H, Lyu X, Liu H (2014) Origin and assessment of groundwater pollution and associated health risk: a case study in an industrial park, northwest China. *Environ Geochem Health* 36(4):693–712
- Lin CY, Abdullah MH, Praveena SM, Yahaya AHB, Musta B (2012) Delineation of temporal variability and governing factors influencing the spatial variability of shallow groundwater chemistry in a tropical sedimentary island. *J Hydrol* 432–433:26–42
- Lo CP, Yeung AKW (2003) Concepts and techniques of geographic information systems. Prentice-Hall of India Pvt. Ltd., New Delhi, p 492
- Loftis JC (1996) Trends in groundwater quality. *Hydrol Process* 10:335–355
- Lopez B, Baran N, Bourguin B (2015) An innovative procedure to assess multi-scale temporal trends in groundwater quality: example of the nitrate in the Seine-Normandy basin, France. *J Hydrol* 522:1–10
- Lu L, Kashiwaya K, Koike K (2016) Geostatistics-based regional characterization of groundwater chemistry in a sedimentary rock area with faulted setting. *Environ Earth Sci* 75:829. <https://doi.org/10.1007/s12665-016-5619-0>
- Lumb A, Sharma TC, Bibeault J-F (2011) A review of genesis and evolution of water quality index (WQI) and some future directions. *Water Qual Exposure Health* 3:1–14
- Machiwal D, Jha MK (2006) Time series analysis of hydrologic data for water resources planning and management: a review. *J Hydrol Hydromech* 54(3):237–257
- Machiwal D, Jha MK (2010) Tools and techniques for water quality interpretation. In: Krantzberg G, Tanik A, Antunes do Carmo JS, Indarto A, Ekdal A (eds) *Advances in water quality control*. Scientific Research Publishing, Inc., USA, pp 211–252
- Machiwal D, Jha MK (2012) *Hydrologic time series analysis: theory and practice*. Springer, the Netherlands and Capital Publishing Company, New Delhi, p 303
- Machiwal D, Jha MK (2014) Role of geographical information system for water quality evaluation. In: Nielson D (ed) *Geographic information systems (GIS): techniques, applications and technologies*. Nova Science Publishers, New York, USA, pp 217–278
- Machiwal D, Jha MK (2015) Identifying sources of groundwater contamination in a hard-rock aquifer system using multivariate statistical analyses and GIS-based geostatistical modeling techniques. *J Hydrol Reg Stud* 4(A):80–110
- Machiwal D, Jha MK, Mal BC (2011) GIS-based assessment and characterization of groundwater quality in a hard-rock hilly terrain of Western India. *Environ Monit Assess* 174:645–663
- Machiwal D, Jha MK, Singh VP, Mohan C (2018) Assessment and mapping of groundwater vulnerability to pollution: current status and challenges. *Earth Sci Rev* 185:901–927. <https://doi.org/10.1016/j.earscirev.2018.08.009>
- MacQueen J (1967) Some methods for classification and analysis of multivariate observations. In: Le Cam L.M, Neyman J (eds). In: *Proceedings of the Fifth Berkeley symposium on mathematical statistics and probability, vol 1*, University of California Press, Berkeley, California, pp 281–297
- Magesh NS, Chandrasekar N, Elango L (2016) Occurrence and distribution of fluoride in the groundwater of the Tamiraparani River basin, South India: a geostatistical modeling approach. *Environ Earth Sci* 75:1483. <https://doi.org/10.1007/s12665-016-6293-y>
- Mair A, El-Kadi AI (2013) Logistic regression modeling to assess groundwater vulnerability to contamination in Hawaii, USA. *J Contam Hydrol* 153:1–23
- Malapati A, Bronson KF, Booker JD, Hudnall WH, Schubert AM (2011) Soil profile sulfate in irrigated Southern High Plains cotton fields and Ogallala aquifer. *J Soil Water Conserv* 66(5):287–294
- Margat J (1968) Groundwater vulnerability maps. Conception-estimation-mapping. EEC Institut Europeen de l' Eau, Paris (**in French**)
- Marko K, Al-Amri NS, Elfeki AMM (2014) Geostatistical analysis using GIS for mapping groundwater quality: case study in the recharge area of Wadi Usfan, western Saudi Arabia. *Arab J Geosci* 7(12):5239–5252
- Masoud AA, Koike K, Mashaly HA, Gergis F (2016) 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. *J Arid Environ* 124:360–376
- Masoud AA, El-Horiny MM, Atwia MG, Gemal KS, Koike K (2018) Assessment of groundwater and soil quality degradation using multivariate and geostatistical analyses, Dakhla Oasis, Egypt. *J Afr Earth Sci* 142:64–81
- Matheron G (1965) *Lee Variables Regionalis ees et leur Estimation*. Masson, Paris, p 306
- Matheron G (1973) The intrinsic random functions and their applications. *Adv Appl Prob* 5:439–468
- McBride GB (2005) *Using statistical methods for water quality management: issues, problems and solutions*. Wiley, New York
- McCuen RH (2003) *Modeling hydrologic change: statistical methods*. Lewis Publishers, CRC Press LLC, Florida
- Melloul A, Collin M (1992) The ‘principal components’ statistical method as a complementary approach to geochemical methods in water quality factor identification; Application to the Coastal Plain aquifer of Israel. *J Hydrol* 140:49–73
- Melloul AJ, Collin M (1998) A proposed index for aquifer water quality assessment: the case of Israel’s Sharon region. *J Environ Manag* 54:131–142

- Mendizabal I, Baggelaar PK, Stuyfzand PJ (2012) Hydrochemical trends for public supply well fields in The Netherlands (1898–2008), natural backgrounds and upscaling to groundwater bodies. *J Hydrol* 450–451:279–292
- Mendoza GA, Martins H (2006) Multi-criteria decision analysis in natural resource management: a critical review of methods and new modelling paradigms. *For Ecol Manag* 230:1–22
- Meng SX, Maynard JB (2001) Use of statistical analysis to formulate conceptual models of geochemical behavior: Water chemical data from the Botucatu aquifer in São Paulo state, Brazil. *J Hydrol* 250:78–97
- Michalak AM, Kitanidis PK (2004) Estimation of historical groundwater contaminant distribution using the adjoint state method applied to geostatistical inverse modeling. *Water Resour Res* 40:W08302. <https://doi.org/10.1029/2004WR003214>
- Mirzaei R, Sakizadeh M (2016) Comparison of interpolation methods for the estimation of groundwater contamination in Andimeshk-Shush Plain, Southwest of Iran. *Environ Sci Pollut Res* 23(3):2758–2769
- Mohammadi K, Niknam R, Majd VJ (2009) Aquifer vulnerability assessment using GIS and fuzzy system: a case study in Tehran-Karaj aquifer. *Iran Environ Geol* 58:437–446
- Molina M, Aburto FN, Calderan RL, Cazanga M, Escudey M (2009) Trace element composition of selected fertilizers used in Chile: phosphorus fertilizers as a source of long-term soil contamination. *Soil Sediment Contam* 18:497–511
- Moore JS (1988) SEEPAGE: a system for early evaluation of pollution potential of agricultural ground water environments. *Geology Technical Note 5 (Revision 1)*, US Department of Agriculture, Soil Conservation Service, Washington, DC
- Mouser PJ, Hession WC, Rizzo DM, Gotelli NJ (2005) Hydrology and geostatistics of a Vermont, USA Kettlehole Peatland. *J Hydrol* 301:250–266
- Myers DE, Begovich CL, Butz TR, Kane VE (1982) Variogram models for regional groundwater geochemical data. *Math Geol* 14(6):629–644
- Nadiri AA, Sedghi Z, Khatibi R, Gharekhani M (2017) Mapping vulnerability of multiple aquifers using multiple models and fuzzy logic to objectively derive model structures. *Sci Total Environ* 593–594:75–90
- Narany TS, Ramli MF, Aris AZ, Sulaiman WNA, Fakharian K (2014) Spatiotemporal variation of groundwater quality using integrated multivariate statistical and geostatistical approaches in Amol-Babol Plain, Iran. *Environ Monit Assess* 186:5797–5815
- Nas B, Berkay A (2010) Groundwater quality mapping in urban groundwater using GIS. *Environ Monit Assess* 160:215–227
- Nasiri F, Maqsood I, Huang G, Fuller N (2007) Water quality index: A fuzzy river-pollution decision support expert system. *J Water Resour Plan Manag ASCE* 133(2):95–105
- National Research Council (1993) Groundwater vulnerability assessment, contaminant potential under conditions of uncertainty. National Academy Press, Washington DC
- Nematollahi MJ, Ebrahimi P, Razmara M, Ghasemi A (2016) Hydrogeochemical investigations and groundwater quality assessment of Torbat-Zaveh plain, Khorasan Razavi, Iran. *Environ Monit Assess* 188:2. <https://doi.org/10.1007/s10661-015-4968-6>
- Niu B, Loáiciga HA, Wang Z, Zhan FB, Hong S (2014) Twenty years of global groundwater research: a science citation index expanded-based bibliometric survey (1993–2012). *J Hydrol* 519:966–975
- Nobre RCM, Rotunno Filho OC, Mansur WJ, Nobre MMM, Cosenza CAN (2007) Groundwater vulnerability and risk mapping using GIS, modeling and a fuzzy logic tool. *J Contam Hydrol* 94:277–292
- Noshadi M, Ghafourian A (2016) Groundwater quality analysis using multivariate statistical techniques (case study: Fars province, Iran). *Environ Monit Assess* 188:419. <https://doi.org/10.1007/s10661-016-5412-2>
- Nriagu JO, Nieboer E (1988) Chromium in the natural and human environments. Wiley-Interscience, New York
- Nshagali BG, Nouck PN, Meli'i JL, Arétouyap Z, Manguelle-Dicoum E (2015) High iron concentration and pH change detected using statistics and geostatistics in crystalline basement equatorial region. *Environ Earth Sci* 73(11):7135–7145
- Ochsenkühn M, Kontoyannakos J, Ochsenkühn-Petropulu M (1997) A new approach to a hydrochemical study of groundwater flow. *J Hydrol* 194:64–75
- Oleson SG, Carr JR (1990) Correspondence analysis of water quality data: Implications for fauna deaths at Stillwater Lakes, Nevada. *Math Geol* 22:665–698
- Otto M (1998) Multivariate methods. In: Kellner R, Mermet JM, Otto M, Widmer HM (eds) Analytical chemistry. Wiley-VCH, Weinheim, p 916
- Pacheco FAL (1998) Finding the number of natural clusters in groundwater data sets using the concept of equivalence class. *Comput Geosci* 24(1):7–15
- Pacheco FAL, Pires LMGR, Santos RMB, Sanches Fernandes LF (2015) Factor weighting in DRASTIC modeling. *Sci Total Environ* 505:474–486
- Panagopoulos G, Antonakos A, Lambrakis N (2006) Optimization of the DRASTIC method for groundwater vulnerability assessment via the use of simple statistical methods and GIS. *Hydrogeol J* 14:894–911
- Paradis D, Vigneault H, Lefebvre R, Savard MM, Ballard J-M, Qian B (2016) Groundwater nitrate concentration evolution under climate change and agricultural adaptation scenarios: Prince Edward Island, Canada. *Earth Syst Dyn* 7(1):183–202
- Park S-C, Yun S-T, Chae G-T, Yoo I-S, Shin K-S, Heo C-H, Lee S-K (2005) Regional hydrochemical study on salinization of coastal aquifers, western coastal area of South Korea. *J Hydrol* 313:182–194
- Parkhurst DL, Thorstenson DC, Plummer LN (1980) PHREEQE: a computer program for geochemical calculations. *Water resources investigations report 80–96*, United States Geological Survey, NTIS Technical Report, PB81-167801, Springfield, Virginia 22161, p 210
- Parkhurst DL, Plummer LN, Thorstenson DC (1982) BALANCE: a computer program for calculating mass transfer for geochemical reactions in ground water. *Water resources investigations Report 82–14*, United States Geological Survey, NTIS Technical Report, PB82-255902, Springfield, Virginia 22161, p 27
- Passarella G, Vurro M, D'Agostino V, Giuliano G, Barcelona MJ (2002) A probabilistic methodology to assess the risk of groundwater quality degradation. *Environ Monit Assess* 79:57–74
- Pebesma EJ, de Kwaadsteniet JW (1997) Mapping groundwater quality in the Netherlands. *J Hydrol* 200:364–386
- Pételet-Giraud E, Dörfliger N, Crochet P (2000) RISKE: method d'évaluation multicritère de la vulnérabilité des aquifères karstiques. Application aux systèmes des Fontanilles et Cent-Fonts (Hérault, Sud de la France) [Risk: methodology for multicriteria evaluation of the vulnerability of karst aquifers. Application to systems Fontanilles and Cent-Fonts Fontanilles (Hérault, southern France)]. *Hydrogéologie* 4:71–88
- Petrișor A-I, Ianoș I, Iurea D, Văidianu MN (2012) Applications of principal component analysis integrated with GIS. *Proc Environ Sci* 14:247–256
- Pettyjohn WA, Savoca M, Self D (1991) Regional assessment of aquifer vulnerability and sensitivity in the conterminous United States. Report EPA-600/2–91/043, United States Environmental Protection Agency, Ada, Oklahoma
- Piper AM (1944) A graphical procedure in the geochemical interpretation of water analysis. *Am Geophys Union Trans* 25:914–928

- Pique A, Grandia F, Canals A (2010) Processes releasing arsenic to groundwater in the Caldes de Malavella geothermal area, NE Spain. *Water Res* 44:5618–5630
- Pirkle FL, Howell JA, Wecksung GW, Duran BS, Stablein NK (1984) An example of cluster analysis applied to a large geologic data set: aerial radiometric data from Copper Mountain, Wyoming. *Math Geol* 16(5):479–498
- Plummer LN, Prestemon EC, Parkhurst DL (1991) An Interactive Code (NETPATH) for Modeling Net Geochemical Reactions along a Flow Path. Water-Resources Investigations Report 94-4078, United States Geological Survey, Reston, Virginia 22092, p 227
- Postma D, Larsen F, Thai NT, Pham TKT, Jakobsen R, Nhan PQ, Long TV, Viet PH, Murray AS (2012) Groundwater arsenic concentrations in Vietnam controlled by sediment age. *Nat Geosci* 5(9):656–661
- Prati L, Pavanello R, Pesarin F (1971) Assessment of surface water quality by a single index of pollution. *Water Res* 5:741–751
- Rainwater FH, Thatcher LL (1960) Methods for Collection and Analysis of Water Samples. 1st edition, United States Geological Survey Water-Supply Paper 1454, Washington, DC
- Ramakrishnaiah CR, Sadashivaiah C, Ranganna G (2009) Assessment of water quality index for the groundwater in Tumkur taluk, Karnataka state, India. *E J Chem* 6(2):523–530
- Ramesh S, Sumukaran N, Murugesan AG, Rajan MP (2010) An innovative approach of drinking water quality index—a case study from southern Tamil Nadu, India. *Ecol Ind* 10:857–868
- Ramos Leal JA, Barrón Romero LE, Sandoval Montes I (2004) Combined use of aquifer contamination risk maps and contamination indexes in the design of water quality monitoring networks in Mexico. *Geofísica Internacional* 43(4):641–650
- Rao SN, Rao SP, Varma D (2013) Spatial variations of groundwater vulnerability using cluster analysis. *J Geol Soc India* 81:685–697
- Rautman CA, Istok JD (1996) Probabilistic assessment of groundwater contamination: 1. Geostatistical framework. *Ground Water* 34(5):899–909
- Reghunath R, Murthy TRS, Raghavan BR (2002) The utility of multivariate statistical techniques in hydrogeochemical studies: an example from Karnataka, India. *Water Res* 36:2437–2442
- Ribeiro L, Macedo ME (1995) Application of multivariate statistics, trend- and cluster analysis to groundwater quality in the Tejo and Sado aquifer. In: Proceedings of the Prague conference on groundwater quality: remediation and protection. IAHS Publication No. 225, Prague, pp 39–47
- Riley JA, Steinhorst RK, Winter GV, Williams RE (1990) Statistical analysis of the hydrochemistry of ground waters in Columbia River basalts. *J Hydrol* 119:245–262
- Rouhani S, Hall TJ (1988) Geostatistical schemes for groundwater sampling. *J Hydrol* 103:85–102
- Rupert MG (2001) Calibration of the DRASTIC ground water vulnerability mapping method. *Ground Water* 39(4):625–630
- Sacha L, Fleming D, Wysocki H (1987) Survey of pesticides in selected areas having vulnerable ground waters in Washington State. EPA/910/9-87/169, United States Environmental Protection Agency, Region X, Seattle, Washington
- Sadat-Noori SM, Ebrahimi K, Liaghat AM (2014) Groundwater quality assessment using the water quality index and GIS in Saveh-Nobaran aquifer, Iran. *Environ Earth Sci* 71:3827–3843
- Saeedi M, Abessi O, Sharifi F, Meraji H (2010) Development of groundwater quality index. *Environ Monit Assess* 163:327–335
- Salas JD, Delleur JW, Yevjevich V, Lane WL (1980) Applied modeling of hydrologic time series. Water Resources Publications, Littleton
- Salah A, Al-Ruwaih F, Shehata M (1999) Hydrogeochemical processes operating within the main aquifers of Kuwait. *J Arid Environ* 42:195–209
- Salman AS, Zaidi FK, Hussein MT (2015) Evaluation of groundwater quality in northern Saudi Arabia using multivariate analysis and stochastic statistics. *Environ Earth Sci* 74(12):7769–7782
- Sánchez-Martos F, Jiménez-Espinoza R, Pulido-Bosch A (2001) Mapping groundwater quality variables using PCA and geostatistics: a case study of Bajo Andarax, southeastern Spain. *Hydrol Sci J* 46(2):227–242
- Sara MN, Gibbons R (1991) Organization and analysis of water quality data. In: Nielsen DM (ed) Practical handbook of ground-water monitoring. Lewis Publishers, Michigan, pp 541–588
- Scanlon BR, Reedy RC, Bronson KF (2008) Impacts of landuse change on nitrogen cycling archived in semiarid unsaturated zone nitrate profiles, Southern High Plains, Texas. *Environ Sci Technol* 42(20):7566–7572
- Scanlon BR, Reedy RC, Gates JB (2010) Effects of irrigated agroecosystems: 1. Quantity of soil water and groundwater in the Southern High Plains, Texas. *Water Resour Res*. <https://doi.org/10.1029/2009WR008427>
- Schaefer JA, Mayor SJ (2007) Geostatistics reveal the scale of habitat selection. *Ecol Model* 209(2–4):401–406
- Schoeller H (1955) Geochimie des eaux souterraines. *Rev de l'Inst Francais du Petrole Paris* 10(3):181–213
- Sener E, Davraz A (2013) Assessment of groundwater vulnerability based on a modified DRASTIC model, GIS and an analytic hierarchy process (AHP) method: The case of Egirdir Lake basin (Isparta, Turkey). *Hydrogeol J* 21:701–714
- Şener Ş, Şener E, Davraz A (2017) Evaluation of water quality using water quality index (WQI) method and GIS in Aksu River (SW-Turkey). *Sci Total Environ* 584–585:131–144
- Sethy SN, Syed TH, Kumar A (2017) Evaluation of groundwater quality in parts of the Southern Gangetic Plain using water quality indices. *Environ Earth Sci* 76:116. <https://doi.org/10.1007/s12665-017-6434-y>
- Shahin M, Van Oorschot HJL, De Lange SJ (1993) Statistical analysis in water resources engineering. A.A. Balkema, Rotterdam
- Shirazi SM, Imran HM, Akib S (2012) GIS-based DRASTIC method for groundwater vulnerability assessment: a review. *J Risk Res* 15(8):991–1011
- Siebert S, Burke J, Faures J, Frenken K, Hoogeveen J, Döll P, Portmann T (2010) Groundwater use for irrigation—a global inventory. *Hydrol Earth Syst Sci* 14:1863–1880
- Skidmore AK, Bijer W, Schmidt K, Kumar L, K (1997) Use of remote sensing and GIS for sustainable land management. *ITC J* 3(4):302–315
- Smith DG (1990) A better water quality indexing system for rivers and streams. *Water Res* 24(10):1237–1244
- Sneath PHA, Sokal RR (1973) Numerical taxonomy. W.H. Freeman and Co., San Francisco
- Snodgrass MF, Kitanidis PK (1997) A geostatistical approach to contaminant source identification. *Water Resour Res* 33(4):537–546
- Soltan ME (1999) Evaluation of groundwater quality in Dakhla Oasis (Egyptian Western Desert). *Environ Monit Assess* 57:157–168
- Sorichetta A, Masetti M, Ballabio C, Sterlacchini S (2012) Aquifer nitrate vulnerability assessment using positive and negative weights of evidence methods, Milan Italy. *Comput Geosci* 48:199–210
- Sprague LA, Lorenz DL (2009) Regional nutrient trends in streams and rivers of the United States, 1993–2003. *Environ Sci Technol* 43(10):3430–3435
- Stafford DB (ed) (1991) Civil engineering applications of remote sensing and geographic information systems. ASCE, New York
- Stamatis G, Parpodis K, Filintas A, Zagana E (2011) Groundwater quality, nitrate pollution and irrigation environmental management in the Neogene sediments of an agricultural region in central Thessaly (Greece). *Environ Earth Sci* 64:1081–1105

- Stambuk-Giljanovic N (1999) Water quality evaluation by index in Dalmatia. *Water Res* 33(16):3423–3440
- Steenhuis TS, Pacenka S, Porter KS (1987) MOUSE: a management model for evaluation ground water contamination from diffuse surface sources aided by computer graphics. *Appl Agric Res* 2:277–289
- Steinhorst RK, Williams RE (1985) Discrimination of groundwater sources using cluster analysis, MANOVA, canonical analysis and discriminant analysis. *Water Resour Res* 21(8):1149–1156
- Steube C, Richter S, Griebler C (2009) First attempts towards an integrative concept for the ecological assessment of groundwater ecosystems. *Hydrogeol J* 17(1):23–35
- Stiff HA Jr. (1951) The interpretation of chemical water analysis by means of patterns. *J Pet Technol* 3(10):15–17
- Stigter TY, Ribeiro L, Carvalho Dill AMM (2006) Application of a groundwater quality index as an assessment and communication tool in agroenvironmental policies: two Portuguese case studies. *J Hydrol* 327:578–591
- Stoner JD (1978) Water-Quality Indices for Specific Water Uses. Geological Survey Circular 770, Washington, DC
- Stumpp C, Żurek A, Wachniew P, Gargini A, Gemtzi A, Filippini M, Witczak S (2016) A decision tree tool supporting the assessment of groundwater vulnerability. *Environ Earth Sci* 75:1057. <https://doi.org/10.1007/s12665-016-5859-z>
- Subbarao C, Subbarao NV, Chandu SN (1996) Characterization of groundwater contamination using factor analysis. *Environ Geol* 28(4):175–180
- Suk H, Lee KK (1999) Characterization of a ground water hydrochemical system through multivariate analysis: clustering into ground water zones. *Ground Water* 37(3):358–366
- Sullivan T, Gao Y (2017) Development of a new P3 (Probability, Protection, and Precipitation) method for vulnerability, hazard, and risk intensity index assessments in karst watersheds. *J Hydrol* 549:428–451
- Sun AY (2007) A robust geostatistical approach to contaminant source identification. *Water Resour Res* 43:W02418. <https://doi.org/10.1029/2006WR005106>
- Sutadian AD, Muttill N, Yilmaz AG, Perera BJC (2016) Development of river water quality indices—a review. *Environ Monit Assess* 188:58. <https://doi.org/10.1007/s10661-015-5050-0>
- Taylor CH, Loftis JC (1989) Testing for trend in lake and ground water quality time series. *J Am Water Resour Assoc* 25(4):715–726
- Tennant CB, White ML (1959) Study of the distribution of some geochemical data. *Econ Geol* 54:1281–1290
- Teso RR, Younglove T, Peterson MR, Sheeks DL, Gallavan RE (1988) Soil taxonomy and surveys: classification of areal sensitivity to pesticide contamination of ground water. *J Soil Water Conserv* 43(4):348–352
- Thirumalaivasan D, Karmegam M, Venugopal K (2003) AHP-DRASTIC: software for specific aquifer vulnerability assessment using DRASTIC model and GIS. *Environ Model Softw* 18(7):645–656
- Thurstone LL (1931) Multiple factor analysis. *Psychol Rev* 38:406–427
- Thyne G, Güler C, Poeter E (2004) Sequential analysis of hydrochemical data for watershed characterization. *Ground Water* 42(5):711–723
- Tryon RC (1939) Cluster analysis. Edwards Brothers, Ann Arbor
- United Nations Environment Programme (2007) Global drinking water quality index development and sensitivity analysis report. United Nations Environment Programme, Global Environment Monitoring System/Water Programme
- USEPA (1996) Guidance for data quality assessment: practical methods for data analysis. Quality Assurance Division, EPA QA/G-9, version QA96, United States Environmental Protection Agency (USEPA), Washington, DC
- USEPA (1998) Guidance for Data quality assessment: practical methods for data analysis. Quality Assurance Division, EPA QA/G-9, version QA97, United States Environmental Protection Agency (USEPA), Washington, DC
- USSL (1954) Diagnosis and improvement of saline and alkaline soils. United States Salinity Laboratory (USSL). In: Richards LA (ed) Hand Book 60. United States Department of Agriculture (USDA), Washington DC, USA, p 159
- Usunoff EJ, Guzman-Guzman A (1989) Multivariate analysis in hydrochemistry: an example of the use of factor and correspondence analysis. *Ground Water* 27(1):27–34
- Vadiati M, Asghari-Moghaddam A, Nakhaei M, Adamowski J, Akbarzadeh AH (2016) A fuzzy-logic based decision-making approach for identification of groundwater quality based on groundwater quality indices. *J Environ Manag* 184:255–270
- Van Stempvoort D, Evert L, Wassenaar L (1992) Aquifer Vulnerability Index: a GIS compatible method for groundwater vulnerability mapping. *Can Water Resour J* 18:25–37
- Vasanthavigar M, Srinivasamoorthy K, Vijayaragavan K, Rajiv Ganthi R, Chidambaram S, Anandhan P, Manivannan R, Vasudevan S (2010) Application of water quality index for groundwater quality assessment: Thirumanimuttar sub-basin, Tamilnadu, India. *Environ Monit Assess* 171(1–4):595–609
- Venkatramanan S, Chung SY, Kim TH, Kim B-W, Selvam S (2016) Geostatistical techniques to evaluate groundwater contamination and its sources in Miryang City. *Korea Environ Earth Sci* 75:994. <https://doi.org/10.1007/s12665-016-5813-0>
- Visser A, Dubus I, Broers HP, Brouyère S, Korcz M, Orban P, Goderniaux P, Batlle-Aguilar J, Surdyk N, Amraoui N, Job H, Pinault JL, Bierkens M (2009) Comparison of methods for the detection and extrapolation of trends in groundwater quality. *J Environ Monit* 11:2030–2043
- von der Heide C, Böttcher J, Deurer M, Weymann D, Duijnsveld WHM (2008) Spatial variability of N₂O concentrations and of denitrification-related factors in the surficial groundwater of a catchment in Northern Germany. *J Hydrol* 360:230–241
- Voudouris K, Polemio M, Kazakis N, Sifaleras A (2010) An agricultural decision support system for optimal land use regarding groundwater vulnerability. *Int J Inf Syst Soc Change* 1(4):66–79
- Vrba J, Zaporozec A (1994) Guidebook on mapping groundwater vulnerability. International contributions to hydrogeology, vol 16. International Association of Hydrogeologists, Hannover
- Wachniew P, Żurek AJ, Stumpp C, Gemtzi A, Gargini A, Filippini M, Rozanski K, Meeks J, Kværner J, Witczak S (2016) Toward operational methods for the assessment of intrinsic groundwater vulnerability: a review. *Crit Rev Environ Sci Technol* 46(9):827–884
- Wagenet RJ, Hutson JL (1987) LEACHM: a finite-difference model for simulating water, salt, and pesticide movement in the Plant root zone, continuum 2. New York State Resources Institute, Cornell University, Ithaca
- WHO (2017) Guidelines for drinking-water quality: fourth edition incorporating the first addendum. World Health Organization, Geneva (**license: CC BY-NC-SA 3.0 IGO**)
- Wilcox LV (1955) Classification and use of irrigation water. Circular 696. United States Department of Agriculture (USDA), Washington, DC
- Williams RE (1982) Statistical identification of hydraulic connections between the surface of a mountain and internal mineralized sources. *Ground Water* 20(4):466–478
- World Bank (2010) Deep wells and prudence: towards pragmatic action for addressing groundwater overexploitation in India. World Bank Report No. 51676, Washington, DC
- Wunderlin DA, del Pilar DM, Valeria AM, Fabiana PS, Cecilia HA, de los Angeles BM (2001) Pattern recognition techniques for the evaluation of spatial and temporal variations in water quality. A case study: Suquia River Basin (Córdoba-Argentina). *Water Res* 35(12):2881–2894

- Wurl J, Mendez-Rodriguez L, Acosta-Vargas B (2014) Arsenic content in groundwater from the southern part of the San Antonio-El Triunfo mining district, Baja California Sur, Mexico. *J Hydrol* 518:447–459
- WWAP (World Water Assessment Programme) (2009) United Nations World Water Development Report 3, water in a changing world, United Nations Educational, Scientific and Cultural Organization (UNESCO), Paris, 2009. <http://unesdoc.unesco.org/images/0021/002156/215644e.pdf>. Accessed 9 May 2017
- WWAP (World Water Assessment Programme) (2012) United Nations World Water Development Report 4, Managing Water Under Uncertainty and Risk, United Nations Educational, Scientific and Cultural Organization (UNESCO), Paris, 2012. <http://unesdoc.unesco.org/images/0021/002156/215644e.pdf>. Accessed 9 May 2017
- Yazdanpanah N (2016) Spatiotemporal mapping of groundwater quality for irrigation using geostatistical analysis combined with a linear regression method. *Model Earth Syst Environ* 2:18. <https://doi.org/10.1007/s40808-015-0071-9>
- Yevjevich VM (1972) Stochastic processes in hydrology. Water Resources Publications, Fort Collins
- Yu WH, Harvey CM, Harvey CF (2003) Arsenic in groundwater in Bangladesh: a geostatistical and epidemiological framework for evaluating health effects and potential remedies. *Water Resour Res* 39(6):1146. <https://doi.org/10.1029/2002WR001327>
- Yue S, Pilon P, Phinney B, Cavadias G (2002) The influence of autocorrelation on the ability to detect trend in hydrological series. *Hydrol Process* 16:1807–1829
- Zakhem BA, Hafez R (2015) Heavy metal pollution index for groundwater quality assessment in Damascus Oasis, Syria. *Environ Earth Sci* 73(10):6591–6600
- Zaporozec A (1972) Graphical interpretation of water-quality data. *Ground Water* 10(2):32–43
- Zwahlen F (ed) (2004) Vulnerability and risk mapping for the protection of carbonate (Karst) aquifers, final report (COST Action 620). European Commission, Directorate-General XII science, research and development, Report EUR 20912, Brussels, p 297