CrossMark

ORIGINAL ARTICLE

Simulation and analysis of temporal changes of groundwater depth using time series modeling

M. Khorasani[1] · M. Ehteshami[2] · H. Ghadimi[3] · M. Salari[4]

Abstract In the world's water scarce regions, groundwater as an important and strategic resource needs proper assessment. An accurate forecasting needs to be performed in order to make a better identification of fluctuating nature of groundwater levels. In this study, groundwater level fluctuations of Kabudarahang aquifer was synchronized and verified. Investigation was conducted with usage and application of time series models. Groundwater level data during 2003–2014 are used for calibration and analyses were performed using Box-Jenkins models. Residual error analysis and comparison of observed and calculated groundwater levels were performed. Then a prediction model for groundwater level in Kabudarahang aquifer developed. The model was used for predicting the groundwater level during 2014–2017. Model results showed that the groundwater level in this aquifer will endure a 5 m decline for the next three upcoming years.

✉ M. Ehteshami
  Maehtesh@gmail.com

  M. Khorasani
  ee.khorasani@gmail.com

  H. Ghadimi
  h_ghadimi@mehr.sharif.ir

  M. Salari
  Salari.marjan@gmail.com

[1] Department of Environment Engineering, K.N. Toosi University of Technology, Tehran, Iran

[2] Civil and Environment Engineering College, K.N. Toosi University of Technology, P.O.Box 1587-544-16, Tehran, Iran

[3] Civil and Environment Engineering College, Sharif University of Technology, Tehran, Iran

[4] Environmental Engineering Department, Shiraz University, Shiraz, Iran

Introduction

Proper management of water resources needs analysis of different hydrological parameters dynamic and complex nature. Groundwater level analysis requires design and assessment in a proper development program. Water level in aquifer is an important parameter in groundwater hydrology; therefore a precise analysis of its temporal and spatial changes can provide valuable information in the aquifer system behavior (Aflatooni and Mardaneh 2011). Groundwater level assessment is a main information source about potential of hydrologic stress within a groundwater system. Groundwater exploitation and land use data during short term and long term periods are valuable in groundwater system evaluations (Winter et al. 2000; Moon et al. 2004; Ahmadi and Sedghamiz 2007).

Models can act as illustrations of simple to complicated hydrogeological occurrences in groundwater assessment strategy (Ahmadi and Sedghamiz 2007; Salami and Ehteshami 2015; Ehteshami et al. 1991). A complex model simulates actions and reactions within each natural process such as groundwater (Ehteshami et al. 1991; Ehteshami et al. 2013; Ehteshami and Biglarijoo 2014). A sufficient model selection has significant importance with regard to many different aspects of groundwater resource management strategy. Proposed models include Man-Kendal method, t test and Cradock tests (Ehteshami et al. 2016), artificial neural network approach (Polemio and Casarano 2008; Salami and Ehteshami 2016), and time series analysis methods like spectral and correlation analysis (Aflatooni and Mardaneh 2011; Nayak et al. 2006; Chen et al.

2004; Lee et al. 2005), and moving average (Rajmohan et al. 2007; Reghunath et al. 2005). In regions where accurate hydrologic data are inaccessible the groundwater level fluctuations can be predicted using stochastic methods like principle component analysis (PCA) and cross correlation analysis methods (Moustadraf et al. 2008).

Conceptual and physical models are the main tools for representing the hydrological variables and perceiving physical processes in a system (Anderson and Woessner 1992). Groundwater systems not only have a dynamic nature but also change in response to climatic stresses and anthropologic activities (Alley et al. 2002). Human induced stresses on groundwater resources include population growth, increasing agricultural productions and rapid expansion of urban areas (Gunatilaka 2005). The main human intervention in groundwater resources is excessive extraction or overuse of resources, particularly in vast urban areas of several Asian countries like China, India, Indonesia, Nepal, Philippine, Thailand and Vietnam (UNEP 2002, ADB 2007, WEPA 2007). While physical modeling is suitable in data scarce conditions, presence of several factors such as anthropologic effects (human interferences) makes accurate determination of hydrological parameters so difficult, therefore application of time series models in these environments are preferred (Kim et al. 2005). A number of time series methods are used in various groundwater studies in order to analyze groundwater data in response to independent stresses (Ganoulis and Morel-Seytoux 1985; Gurdak et al. 2007; Perry 1994; Perry and Hsu 2000; Yu and Chu 2012).

In order to evaluate the groundwater potential often a systematic approach with ability of trend decomposition, analysis of time series and frequency analysis and comparing between various types of data and frequencies in hydrologic setting and between different basins is required for local and regional assessments (Hanson et al. 2004). A coherent and reasonable decomposition method can separate and quantitatively analyze the effects of independent factors on groundwater and also identify the relations and complexities between such parameters in hydrological processes and climatological changes (Hanson et al. 2004; Mayer and Congdon 2008). Various frequency analysis methods can extract seasonal characteristics. Detrending is the process of eliminating a trend from a time series in order to remove a property that can cause long-term effects on time series such as groundwater extraction and irrigation application of groundwater which can lead to irregularities in some embedded cyclic signals (Hanson et al. 2004; Gardner and Heilweil 2009; Prinos et al. 2002).

There are number of different methods and techniques for analyzing groundwater table fluctuation through probability characteristics, time series methods, synthetic data generation, multiple regression, group theory, pattern recognition and neural network methods (Adhikary et al. 2012). A time series model can be defined as an empirical model for stochastic modeling and forecasting temporal behavior of hydrologic systems (Brockwell and Davis 2002). The stochastic time series models are useful and popular tools for medium-range forecasting and synthetic data generation. Several stochastic time series models including the Markov, Box- Jenkins (BJ) Seasonal Autoregressive Integrated Moving Average (SARIMA), deseasonalized Autoregressive Moving Average (ARMA), Periodic Autoregressive (PAR), Transfer Function Noise (TFN) and Periodic Transfer Function Noise (PTFN), can be used for these purposes (Box et al. 2008). The Markov, ARMA and SARIMA models are univariate models and the PAR, TFN and PTFN models are multivariate models. Furthermore, the PAR and PTFN models are periodic multivariate models (Hipel and Mcleod 1994).

The selection of a proper modelling method for a particular problem depends on many factors such as the number of series that are being modelled, required accuracy, modelling costs, ease of model usage and results interpretation, etc. (Mondal and Wasimi 2007) Several applications of all of these models for analysis of fluctuations of groundwater level over time are vastly used in several groundwater hydrology applications (Mondal and Wasimi 2006). Furthermore, they have extensive use in various areas of science and engineering such as forecasting of river discharge and synthetic data generation (Salas and Obeysekera 1982; Mondal and Wasimi 2003) stream flow and water quality data prediction (Mondal and Wasimi 2005; Kurunc et al. 2005). Water resource monitoring and assessment (Mondal and Wasimi 2003) and drought period simulation (Durdu 2010). When a relatively few number of series need to be modeled and a large expenditure of time and effort can be justified using the BJ Method (Seasonal ARIMA) is of priority (Mondal and Wasimi 2007). The purpose of this study is to develop SARIMA models for application in groundwater level fluctuation modeling by developing equations through usage of seasonal ARIMA algorithms.

## Materials and methods

Study area

Kabudarahang as one of the main aquifers in Northern part of Hamedan County covers 3448 Km^2 of Gharachay river basin. The Kabudarahang County with an average altitude of 1680 m above mean sea level is located at approximately $35°12'30''$ north latitude, $48°43'26''$ east longitude within a 60 km distance from the capital city of the Hamdan province, Iran with an area of 1186 Km^2. The

mean annual regional precipitation for a recorded 36 years is 332.7 mm (Fig. 1).

Time series modeling

Time series modeling or in other words describing the behavior of a time series in mathematics language includes three common steps: identifying the experimental model, estimating the model parameters (fitting), and verification of the model. In order to proper identification of an experimental model, it is recommended that at least 50 observations of the regarded series must be available. In analyzing hydrological and environmental time series with application of stochastic and statistical theories it is assumed that all variables have a normal distribution. First step in modeling a time series is to draw it as a time series
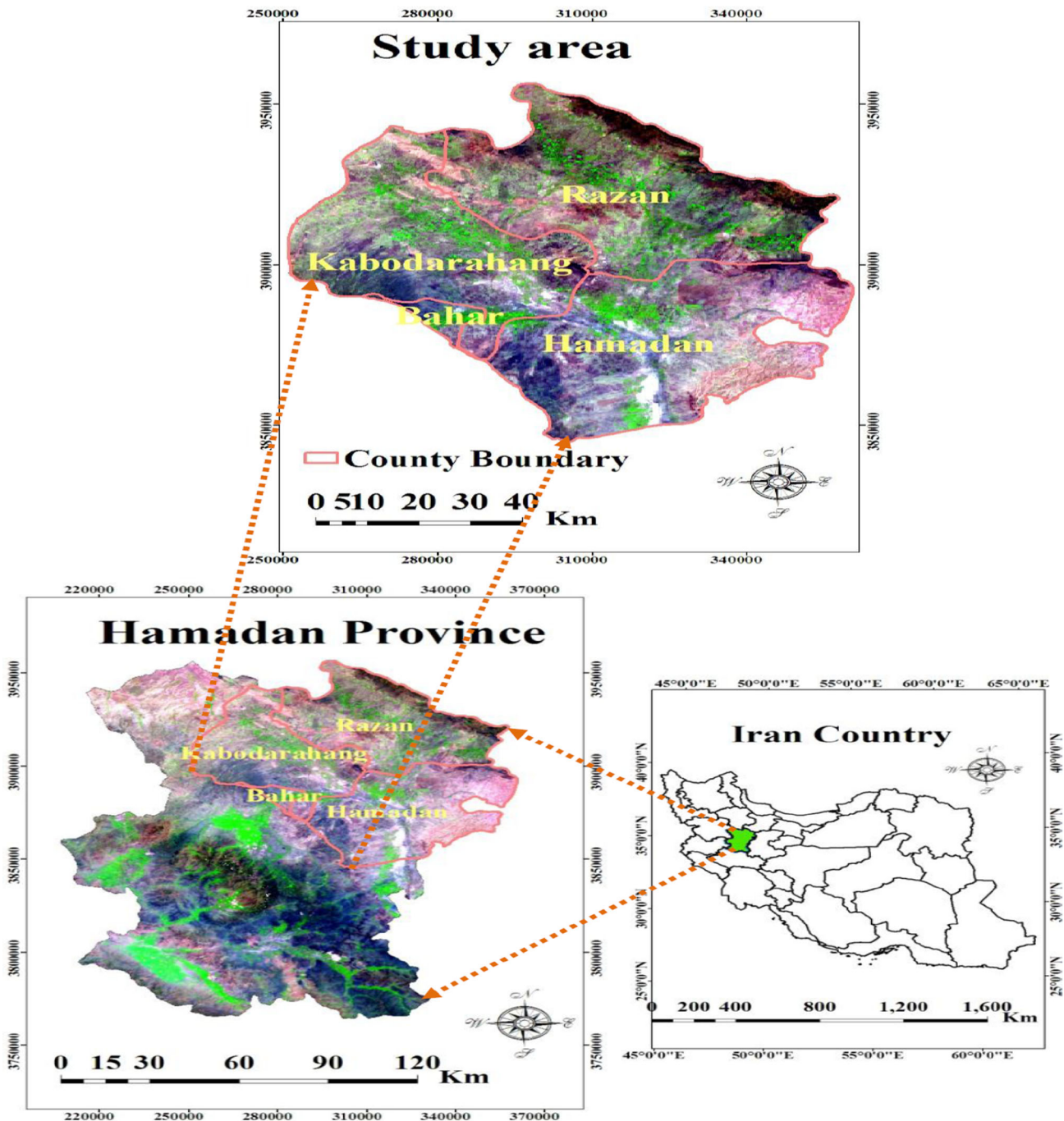
Fig. 1 The status of study region

diagram which can give a boost for identification of trends, variance non-stationarity, seasonality (periodicity) and other types of irregularities in data.

Most of time series data doesn't follow a normal distribution, thus before any modeling or analyzing processes they need to be normalized by means of transfer functions (Fig. 2). As a pretreatment step, a variety of different transfer functions and statistical tests are available to prepare data before any modeling which includes Box- Cox transformation and non-stationarity in mean or variance. These tests can be categorized in two main groups: independence tests (for time), and normal distribution tests. Also prior to any modeling process the conditions of time series are tested for goodness of fit, well-fitting criteria, level of reliability and statistical parameters limits.

Tests of independence

For modeling the irregular (random noise) component, the time series should be time independent, which means each value should be independent from its previous and subsequent values. For instance, after removing the trend, seasonality and shock components, the mean annual precipitation should be independent of previous and the next year mean precipitation in order to be time independent. Correlogram (autocorrelation function) is a useful device for describing the behavior of stationary processes which can determine if they are independent of time. In

fact the autocorrelation function shows the variations of autocorrelation coefficients through different time lags. The autocorrelation coefficient at lag K that is the correlation between X_t and X_{t+k} can be shown by r_k which is:

$$r_k = \frac{C_k}{C_0} = \frac{\sum_{t=1}^{N-K}(X_t - \bar{X})(X_{t+k} - \bar{X})}{\sum_{t=1}^{N}(X_t - \bar{X})^2}, \qquad (1)$$

where $C_k (= E[(X_{t+k} - \bar{X})(X_t - \bar{X})])$ is defined as autocovariance function and therefore C_0 also can be defined as the variance of the time series. The Anderson method, turning point method and Porte Manteau method are three principal methods for analyzing time independency of time series and to diagnose independency of time in periodic series the cumulative period gram method can be used.

Anderson test of independence in time

In this method confidence limits of 95 and 99 % for the autocorrelation coefficient are defined as:

$$r\_k\,(95\%) = (-1 \pm 1.645\sqrt{(N - k - 1)})/(N - k), \qquad (2)$$

$$r_k\,(99\%) = \frac{-1 \pm 2.326\sqrt{N - k - 1}}{N - k}, \qquad (3)$$

where N is the number of the sample and k is the lag. The autocorrelation functions within these limits are independent of time and else they have time dependency. It is noteworthy to mention that if for each value of k, $r_k = 0$ it implies that the time series is quite independent of time or serially independence (Yurekli et al. 2004).

Ests of normality

There are several methods to check the normality of time series, one of which is a common method that is to plot the time series in a normal probability plot and checking if the plots form a straight line. If they form a rather straight line it implies that the time series has a normal distribution, otherwise the normality assumption is refuted. In addition to this method, the two tests of Chi square test and skewness test are applicable for examining normal distribution of time series. The normal distribution of groundwater data is tested by Chi square test as follows.

Chi square test

This method is preferably used for clustered data. At first the normal distribution function is fitted to the sample time series, and then the goodness of fit is quantitatively evaluated as probability percentile. When X_t denote the sample time series with the mean $\bar{X}$ and variance $\hat{\sigma}$. This series can be classified into k groups with the attributed probability of $\frac{1}{k}$ for each in an ascending manner. The values of
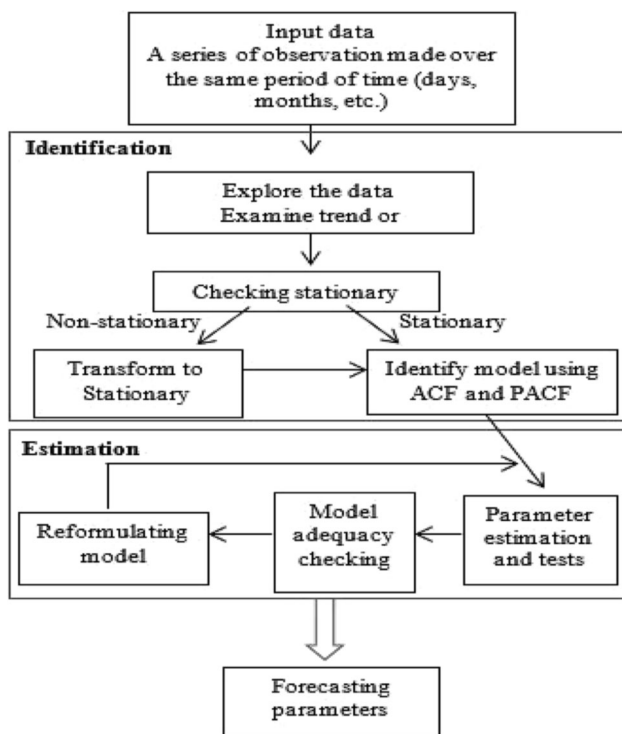

Fig. 2 Functional diagram of the Box-Jenkins modeling strategy

$u_1, u_2, u_3, \ldots, u_{k-1}$ for the cumulative probability of $\frac{1}{k}, \frac{2}{k}, \frac{3}{k}, \ldots, \frac{k-1}{k}$ can be obtained from the normal distribution table. Therefore the following values can be calculated using the obtained values from the table and the mean and variance of the sample time series such as:

$$\overset{\prime}{X}_1 = \bar{X} + \hat{\sigma} u_1, \tag{4}$$

$$\overset{\prime}{X}_2 = \bar{X} + \hat{\sigma} u_2, \tag{5}$$

$$\overset{\prime}{X}_3 = \bar{X} + \hat{\sigma} u_3, \ldots, \overset{\prime}{X}_{k-1} = \bar{X} + \hat{\sigma} u_{k-1}. \tag{6}$$

Let N_i denote absolute frequency of the ordered sample series for group i. Therefore the number of points which are expected to set in each group is equal to $\frac{N}{k}$. Now Chi square test with freedom degree of (k-2) is applied considering the confidence level of α. Finally the normality of the sample time series can be evaluated by comparing the results with the following value:

$$\chi^2 = \sum_{i=1}^{k} \frac{\left(N_i - \frac{N}{k}\right)^2}{\left(\frac{N}{k}\right)}. \tag{7}$$

In this case, if the value $\chi^2_{1-\alpha}(K-2)$ obtained from the Chi square table is greater than χ^2 from the formula, the normal assumption for distribution of time series X_t is true, otherwise the assumption of normal distribution is refuted (Karamouz et al. 2012). Summary of Properties of autoregressive (AR), moving average (MA), and mixed autoregressive moving average (ARMA) processes show that in Table 1.

Drawing graphs of the ACF and PACF

The autocorrelation function (ACF) and partial autocorrelation function (PACF) are useful tools for identifying models. The type and order of the process can be determined by means of graphical analysis using these diagrams. In practice, to obtain a useful estimate of the autocorrelation function, a minimum sample population (n = number of observations) of 50 is recommended (Mondal and Wasimi 2003) and the ACFs and PACFs should be calculated and depicted for k = 1, 2, 3, …, K where K is not larger than $\frac{n}{4}$, for a more reliable identification of the model. For a given observed time series of $Z_1, Z_2, \ldots, Z_n$, the sample ACF is defined by:

$$\rho_k = \frac{\hat{\gamma}_k}{\hat{\gamma}_0} = \frac{\sum_{t=1}^{n-k}(Z_t - \bar{Z})(Z_{t+k} - \bar{Z})}{\sum_{t=1}^{n}(Z_t - \bar{Z})}. \tag{8}$$

Hence, the graph in which $\hat{\rho}_k$ is plotted against k is known as the sample correlogram. To calculate the sample PACF ($\hat{\phi}_{kk}$), (Chandler and Scott 2011) represents following equations:

Table 1 Summary of properties of autoregressive (AR), moving average (MA), and mixed autoregressive moving average (ARMA) processes

	AR (p)	MA (q)	ARMA (p:q)
Model	$w_t = \phi_1 w_{t-1} + \cdots + \phi_p w_{t-p} + a_t$	$w_t = a_t - \theta_1 a_{t-1} - \cdots - \theta_q a_{t-q}$	$w_t = \phi_1 w_{t-1} + \cdots + \phi_p w_{t-p} - \theta_1 a_{t-1} + \cdots - \theta_q a_{t-q} + a_t$
Autocorrelation Function (ACF)	Infinite: damped Exponentials and/o damped since waves: Tails off	Finite: cuts off after q lags	Infinite: damped Exponentials and/or damped since waves: Tails off
Partial autocorrelation Function (PACF)	Finite: cuts off after p lags	Infinite: damped Exponentials and/or damped since waves: Tails off	Infinite: damped Exponentials and/or damped since waves: Tails off

$$phi_{k+1,k+1} = \frac{\hat{\rho}_{k+1} - \sum_{j=1}^{k} \hat{\phi}_{kj}\hat{\rho}_{k+1-j}}{1 - \sum_{j=1}^{k} \hat{\phi}_{kj}\hat{\rho}_{j}}, \tag{9}$$

and:

$$\hat{\phi}_{k+1,j} = \hat{\phi}_{kj} - \hat{\phi}_{k+1,k+1}\hat{\phi}_{k,k+1-j}. \tag{10}$$

The ACF and PACF diagrams of groundwater level fluctuation time series of Kabudarahang plain are shown in Fig. 3.

Model selection

Selecting the most reliable modeling technique between several available models for seasonal forecasting and analyzing requires an inductive approach. The techniques are such as: (1) exponential smoothing; (2) Markov models; (3) decomposition and trend extrapolation method; (4) Holt–Winters method; and (5) Box-Jenkins Seasonal Autoregressive Integrated Moving Average (SARIMA). The model selection depends on many factors such as: number of series to be forecasted, required accuracy of forecasts, model facility and proficiency, modeling costs, interpretative outputs, etc.

Decomposition method

An alternative way of modeling time series data is based on the decomposition of the model into trend, seasonal, and noise components. The classical decomposition partitions a signal into three elemental components called trend, periodicity and random or irregular components and can be written in multiplicative form as

$$x_t = m_t \times s_t \times y_t. \tag{11}$$

X is the time series data, m_t is the trend component, s_t is the periodic component, and y_t is random noise or irregular component. There are a number of techniques available to evaluate trends within datasets. The simplest way is to apply a simple linear regression analysis or some low-order polynomial regression (Durbin 1960). Moving average, e.g. a 12-month moving average, is an alternative detrending
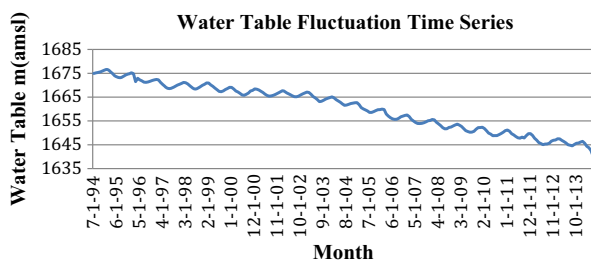

Fig. 3 Time series of observed groundwater level fluctuations in Kabudarahang aquifer

method to remove small-scale structures, e.g. noises, and short periods from a time series.

In extrapolation of trend curves method for decomposition of time series, a trend line is fitted to the smoothed data using the least squares regression. Then the series become detrended by either dividing the data by the trend component (multiplicative model) or subtracting the trend component from the data (additive model). However, Minitab trend analysis tool uses the multiplicative model by default. Subsequently a centered moving average with a length equal to the length of the seasonal cycle is applied to smooth the detrended data. At the next step the raw seasonal values are calculated by either dividing the moving average into multiplicative model or subtracts it from the additive model. The median of the raw seasonal values for corresponding periods of time in the seasonal cycles are determined and adjusted to constitute the seasonal indices which are used to seasonally adjust the data. Therefore, the seasonal adjustment procedure is applied at this step by the iterative application of moving averages according to sequence of the seasonal adjustment procedure (Chatfield 2011).

Seasonal ARIMA model: SARIMA(p,d,q) × (P,D,Q)s

ARIMA models generally have both non-seasonal and seasonal components. The general non-seasonal ARIMA model is AR to order p and MA to order q and operates on the dth difference of the time series z_t; thus, a model of the ARIMA family is classified by three parameters (p, d, q) that can have zero or positive integral values. have generalized the ARIMA model to deal with seasonality and have defined a general multiplicative seasonal ARIMA model, commonly known as a seasonal ARIMA model. In this study the general mixed seasonal model is denoted as: ARIMA(P,D,Q)s. Where, P = order of the seasonal autoregressive process, D = number of seasonal differencing, Q = order of the seasonal moving average and s = the number of the seasonality. The seasonal ARIMA model described as ARIMA $(p, d, q).(P, D, Q)S$, where (p, d, q) non-seasonal part of the model and (P, D, Q) seasonal part of the model with a seasonality S, can be written as:

$$\phi_p(B)Phi_p(B^s)\nabla^d\nabla^D_S z_t = \theta_q(B)\Theta_Q(B^s)a_t. \tag{12}$$

where, $\phi(B)$ and $\phi(B)$ are polynomials of order p and q, respectively; $\Phi(B^s)$ and $\Theta(B^s)$ are the polynomials in B^s of degrees P and Q respectively; p equals order of non-seasonal autoregressive operator; d equals, number of regular differencing; q order of the non-seasonal moving average; P order of seasonal autoregression; Q order of seasonal moving average. The ordinary and seasonal difference components are designated by:

$$\nabla^d = (1 - B)^d \quad \text{and} \quad \nabla_s^D = (1 - B^s)^D. \tag{13}$$

In derivations of d and D; B is the backward shift operator, d is the number of regular differences, D is the number of seasonal differences, S = seasonality. Zt denotes observed value at time t, where $t = 1, 2\dots k$; and at is the Gaussian white noise or estimated residual at time t.

Holt–Winters forecasting method

Exponential smoothing can be easily generalized to time series which include trends and seasonal changes. This practice is known as Holt-Winters method and was introduced by Winters and Chatfield to describe trend and seasonal components which are synchronized using exponential smoothing algorithms (Montgomery et al. 2008). By assuming that observed parameters are monthly and m_t represents average estimation, r_t represents trend estimation (expected increase or reduction per month in current average) and S_t represents seasonal operating of month t. In this case since every new observation is possible, all three components are synchronized. If seasonal variations are multiplicative, synchronizing equations would be (Mondal and Wasimi 2005):

$$m_t = \alpha x_t / s_{t-1,2} + (1 - \alpha)(m_{t-1} + r_{t-1}), \tag{14}$$

$$s_t = \beta x_t / m_t + (1 - \beta)s_{t-1,2}, \tag{15}$$

$$r_t = \gamma(m_t - m_{t-1}) + (1 - \gamma)r_{t-1}, \tag{16}$$

where x_t denotes the latest observation and α, β, γ are constants which $0 < \alpha, \beta, \gamma < 1$, so the predictions about time t are:

$$\hat{X}(t, h) = (m_t + hr_t)s_{t-1,2+h}, \quad (h = 1, 2, \dots, 12). \tag{17}$$

If variations are cumulative, seasonal equations would be:

$$m_t = \alpha(x_t - s_{t-1,2}) + (1 - \alpha)(m_{t-1} + r_{t-1}), \tag{18}$$

$$s_t = \beta(x_t - m_t) + (1 - \beta)s_{t-1,2}, \tag{19}$$

$$r_t = \gamma(m_t - m_{t-1}) + (1 - \gamma)r_{t-1}, \tag{20}$$

To see whether the cumulative seasonal variable is better or the multiplicative seasonal variable, data diagram should be tested. Therefore, if the time series exhibits an obvious trend or seasonal pattern it indicates that the appropriate model is chosen. Moreover, if the seasonal period doesn't include 12 months observations, the equations should be corrected. Basic values of s_t, r, m_t can be selected in a relatively inexperienced way using data of first two years considering three constants α, β, γ with the difference of minimum quantity equals to $\sum_{i=2,5}^{N} e_i^2$.

Model verification

In order to evaluate the models the following indicators are used:

$$SSE = \sum_{i=1}^{n} (T_{actual} - T_{forecast})^2, \tag{21}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(T_{actual} - T_{forecast})^2}{n}}, \tag{22}$$

$$MAE = \frac{\sum_{i=1}^{n}|T_{actual} - T_{forecast}|}{n}. \tag{23}$$

As it is presented the minimum residuals are belonging to the $SARIMA(1, 1, 0)(1, 1, 1)12$ model which has the best fitting to observed data.

Results and discussion

A comparative study on state-of-the-art prediction tools such as Box-Jenkins and Holt-Winters have been used to evaluate groundwater level fluctuations. Also, the extrapolation of trend curve was performed to smooth the detrended data. The assumption of zero constant term (Constant, SAR, AR) proved to be acceptable considering satisfactory statistical t test and P test values. Therefore the model can be reliably fitted without a constant term in order to represent stochastic trends. For instance, for some SARIMA models P value is greater than 0.05. Thus the assumption of $H_0 : \theta_0 = 0$, $H_1 : \theta_1 = 0$, $H_2 : \theta_2 = 0$ should be accepted with 95 % confidence level, for other models also the zero assumption of constant term and SAR and AR coefficients are reasonable since the modeled differencing series has a nearly zero mean. In order to evaluate the eligibility of the model a residual analysis of the fitted model performed. Residual analysis results of the fitted model in the Kabudarahang plain's hydrographs are shown in Table 2. The normal probability plot and histogram of residual values confirm that the normal assumption can be accepted.

Similarly, these verifications have been applied on other models too. Furthermore, the Porte-Manteau method is used as a more formal method for model verification which is based on autocorrelation of residuals. These results for different models are evaluated. Since the p value, considering different lags for the model, is desirably more than 0.05 the assumption of zero autocorrelations up to lag 48 is acceptable such as, $(H_0 : \rho_1 = \rho_2 = \cdots = \rho_{48} = 0)$.

Finally the Akaike criterion is used for model selection. Considering Akaike standard as:

$$AIC = n \times \ln \hat{\sigma}^2 + 2(K + 1), \quad \hat{\sigma}^2 = \frac{SSE}{n}. \tag{24}$$

Table 2 Model verification: evaluating the goodness of fit

Model	SSE	RMSE	MAE
SARIMA(1,1,0)(1,1,1)12	27.3227	0.346936	0.2145
Holt-winters	67.5346	0.545444	0.40476
Decomposition	388.856	1.308825	1.02995

Table 3 Estimates of fitting ARIMA(1,1,0)(1,1,1)12 model to Kabudarahang plain groundwater level data

Model term	Estimate	Standard error	t	p
AR1	−0.0512	0.0712	−0.72	0.472
SAR12	−0.0344	0.0835	−0.41	0.681
SMA12	0.9128	0.0532	17.17	0.000
Constant	−0.0004352	0.0003580	−1.22	0.225
Lag	12	24	36	48
Chi square	7.4	16.1	26.9	29.7
DF	8	20	32	44
P value	0.496	0.712	0.723	0.951

Differencing: 1 regular, 1 seasonal of order 12

Number of observations: original series 240, after differencing 227

Residuals: SS = 0.297838; MS = 0.001336; df = 223

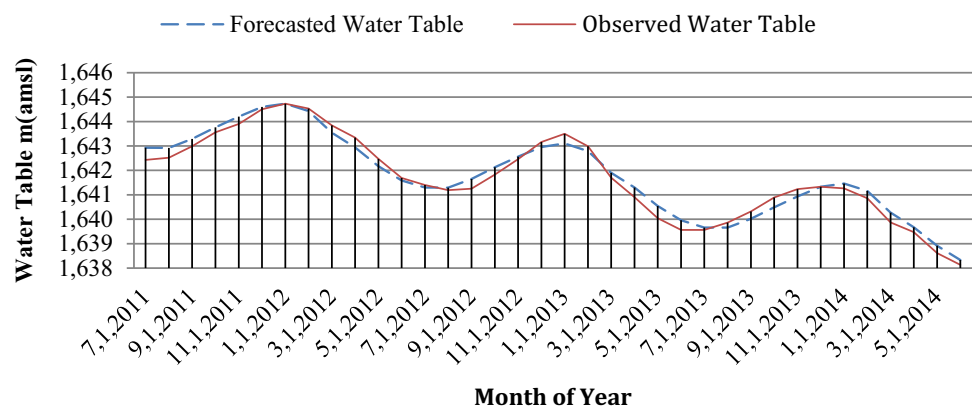Modified box-pierce (Ljung-box) Chi square test

Therefore SARIMA(1,1,0)(1,1,1)12 selected as the best model among other seasonal models of Box-Jenkins for groundwater level simulation due to smaller Akaike value in comparison with other models(Tables 3, 4).

The results of evaluated models and the real values of groundwater level (during 2003–2014) and anticipated values (during 2014–2017), Holt model, extrapolation of trend and Box-Jenkins model for level changes, SARIMA(1,1,0)(1,1,1)12 model reveals more correlation between measured groundwater level fluctuation data and simulated values (Tables 3, 4). Figure 4 shows measured and forecasted groundwater depths. As it is shown, the selected models predict an overall decreasing trend in mean ground water levels which results in a 5 m decline during three upcoming years. The forecasted results and measured data show a good agreement ($R^2 = 98$ %).

Conclusions

This study presents a data conservative approach of time series modeling to evaluate the groundwater fluctuations in semi-arid area of Kabudarahang plain. Several stochastic models were developed for groundwater

Table 4 Akaike information criterion analysis for seasonal ARIMA models

Model	n	SSE	$(\hat{\sigma})^2$	k	AIC value
SARIMA(0,1,1)(1,1,0)12	227	0.297836	0.001312053	2	−1500.41
SARIMA(1,1,0)(1,1,0)12	227	0.366473	0.001614419	2	−1453.33
SARIMA(1,1,1)(1,1,1)12	227	0.297696	0.001311436	4	−1496.52
SARIMA(1,1,1)(1,1,0)12	227	0.298423	0.001314639	3	−1497.96
SARIMA(1,1,0)(0,1,1)12	227	0.298433	0.001314683	2	−1499.95
SARIMA(0,1,1)(0,1,1)12	227	0.297831	0.001312031	2	−1500.41
SARIMA(1,1,1)(0,1,1)12	227	0.297838	0.001312062	3	−1498.41
SARIMA(0,1,1)(1,1,1)12	227	0.000100094	4.40943E−07	3	−3314
SARIMA(1,1,0)(1,1,1)12	227	0.000104857	4.61925E−07	3	−3303.44

Fig. 4 Forecasted groundwater level by the selected SARIMA model

assessment. Then, after data preparation, diagnostic checking, and model verification, by usage of Akaike criteria, adequate model was selected. Consequently, the selected model was used to forecast the groundwater table for the next 3 years. As it is discussed the seasonal ARIMA algorithms proved to be a useful tool for simulation and short-term forecasting with a reasonably high correlation coefficient of 98 % to the measured data considering their dynamic and complex nature of groundwater. The presented results may be useful for regulatory agencies to make the best use of funds available for monitoring groundwater level fluctuations.

References

ADB (2007) Recent advances in water resources development and management in developing countries in Asia. Asian Water Development Outlook. Asian Development Bank, Manila

Adhikary SK, Rahman M, Gupta AD (2012) A stochasticmodelling technique for predicting groundwater table fluctuations with time series analysis. Int J Appl Sci Eng Res 1(2):238–249

Aflatooni M, Mardaneh M (2011) Time series analysis of groundwater tablefluctuations due to temperature and rainfall change in Shiraz plain. Int J Water Resour Environ Eng 3:176–188

Ahmadi SH, Sedghamiz A (2007) Geostatistical analysis of spatial and temporal variations of groundwater level. Environ Monit Assess 129(1–3):277–294

Alley WM, Healy RW, LaBaugh JW, Reilly TE (2002) Flow and storage in groundwater systems. Science 296:1985–1990

Anderson MP, Woessner WW (1992) Applied groundwater modeling: simulation of flow and advective transport. Academic Press, New York

Box GEP, Jenkins GM, Reinsel GC (2008) Time series analysis: forecasting and control, 4th edn. Prentice Hall, Englewood Cliffs

Brockwell PJ, Davis RA (2002) Introduction to time series and forecasting, vol 1. Taylor & Francis, London

Chandler RE, Scott EM (2011) Statistical methods for trend detection and analysis in the environmental sciences. Wiley, Chichester, UK

Chatfield C (2011) The analysis of time series. Wiley, London

Chen Z, Grasby S, Osadetz KG (2004) Relation between climate variability and groundwater levels in the upper carbonate aquifer, southern Manitoba, Canada. J Hydrol 290:243–262

Durbin J (1960) The fitting of time-series models. Rev Inst Int Stat 28(3):233–244

Durdu ÖF (2010) Stochastic approaches for time series forecasting of boron: a case study of Western Turkey. Environ Monit Assess 169(1–4):687–701

Ehteshami M, Biglarijoo N (2014) Determination of nitrate concentration in groundwater in agricultural area in Babol County. Iran J Health Sci. 2(4):1–9

Ehteshami M, Peralta RC, Eisele H, Deer H, Tindall T (1991) Assessing pesticide contamination to ground water: a rapid approach. J Ground water 29(6):862–886

Ehteshami M, Langeroudi AS, Tavassoli S (2013) Simulation of nitrate contamination in groundwater caused by livestock

industry (Case study: Rey). J Environ Protect 4(7):91–97. doi:10.4236/jep.47A011

Ehteshami M, Dolatabadi Farahani N, Tavassoli S (2016) Simulation of nitrate contamination in groundwater using artificial neural networks. J Model Earth Syst Environ (Springer) 2:28. doi:10.1007/s40808-016-0080-3

Ganoulis J, Morel-Seytoux H (1985) Application of stochastic methods to the study of aquifer systems. UNESCO, Technical documents in hydrology, Paris

Gardner PM, Heilweil VM (2009) Evaluation of the effects of precipitation on ground-water levels from wells in selected alluvialaquifers in Utah and Arizona, 1936–2005. US Geological Survey Scientific Investigations Report, Utah and Arizona

Gunatilaka A (2005) Groundwater woes of Asia. Asian Water. January/February, 2005, pp 19–23

Gurdak JJ, Hanson RT, McMahon PB, Bruce BW, McCray JE, Thyne GD (2007) Climate variability controls on unsaturated water and chemical movement, High Plains Aquifer, USA. Vadose Zone J 6(3):533–547

Hanson RT, Newhouse MW, Dettinger MD (2004) A methodology to assess relations between climatic variability and variations in hydrologic time series in the southwestern United States. J Hydrol 287(1):252–269

Hipel KW, Mcleod AI (1994) Time series modelling of water resources and environmental systems. Elsevier Science, Amsterdam

Karamouz M, Nazif S, Falahi M (2012) Hydrology and hydroclimatology: principles and applications. Taylor & Francis, London

Kim SJ, Hyun Y, Lee KK (2005) Time series modeling for evaluation of groundwater discharge rates into an urban subway system. Geosci J 9(1):15–22

Kurunc A, Yurekli K, Cevik O (2005) Performance of two stochastic approaches for forecasting water quality and stream flow data from Yesilirmak River, Turkey. Environ Modell Softw 20(9):1195–1200

Lee J, Choi Y, Kim MJ, Lee KK (2005) Evaluation of hydrologic data obtained from a local groundwater monitoring network in a metropolitan city, Korea. Hydrol Processes 19(13):2525–2537

Mayer TD, Congdon RD (2008) Evaluating climate variability and pumping effects in statistical analyses. Ground Water 46(42):212–227

Mondal M, Wasimi S (2003) Forecasting of seasonal flow of the Ganges River in Bangladesh with SARIMA model. In: Second annual paper meet and international conference on civil engineering, Dhaka

Mondal M, Wasimi S (2005) Periodic transfer function-noise model for forecasting. J Hydrol Eng 10(5):353–362

Mondal M, Wasimi S (2006) Generating and forecasting monthly flows of the Ganges river with PAR model. J Hydrol 323:41–56

Mondal M, Wasimi S (2007) Choice of model type in stochastic river hydrology. In: 1st international conference on water and flood management, 12–14 March 2007. Dhaka, Bangladesh

Montgomery DC, Jennings CL, Kulahci M (2008) Introduction to time series analysis and forecasting, 2nd edn. Wiley, London

Moon SK, Woo NC, Lee KS (2004) Statistical analysis of hydrographs and water-table fluctuation to estimate groundwater recharge. J Hydrol 292:198–209

Moustadraf J, Razack M, Sinan M (2008) Evaluation of the impacts of climate changes on the coastal Chaouia aquifer, Morocco, using numerical modeling. Hydrogeol J 16:1411–1426

Nayak CP, Sayajirao YR, Sudheer KP (2006) Groundwater level forecasting in a shallow aquifer using artificial neural network approach. Water Resour Manage 20:77–90

Perry CA (1994) Solar-irradiance variations and regional precipitation fluctuations in the western United States. Int J Climatol 14:969–983

Perry CA, Hsu KJ (2000) Geophysical, archaeological, and historical evidence support a solar-output model for climate change. Proc Natl Acad Sci 97(23):1244–1248

Polemio M, Casarano D (2008) Climate change, drought and groundwater availability in southern Italy. Geol Soc Spec Publ 288:239–251

Prinos ST, Leitz AC, Irvin RB (2002) Design of a real-time groundwater level monitoring network and portrayal of hydrologic data in southern Florida. US geological survey water-resources investigations repor, Florida, pp 01–4275

Rajmohan N, Al-Futaisi A, Jamrah A (2007) Evaluation of long-term groundwater level data in regular monitoring wells, Barka, Sultanate of Oman. Hydrol Process 21:3367–3379

Reghunath R, Murthy T, Raghavan B (2005) Time series analysis to monitor and assess water resources: a moving average approach. Environ Monit Assess 109(1–3):65–72

Salami ES, Ehteshami M (2015) Simulation evaluation and prediction modeling of river water quality properties (Case study: Ireland Rivers). Int J Eng Sci Technol (Springer) 12(10):3235–3242. doi:10.1007/s13762-015-0800-7

Salami ES, Ehteshami M (2016) Application of artificial neural networks to estimating DO and salinity in San Joaquin River basin. Desalination Water Treat. 57(11):4888–4897. doi:10.1080/19443994.2014.995713

Salas JD, Obeysekera J (1982) ARMA model identification of geophysical time series. Water Resour Res 18:1011–1021

United Nations Environment Programme (UNEP) (2002) Global environment outlook 3. Nairobi, Kenya

WEPA (2007) Water Environment Partnership in Asia (WEPA). http://www.wepa-db.net/index.htm

Winter TC, Mallory SE, Allen TR, Rosenberry DO (2000) The use of principal component analysis for interpreting ground water hydrographs. Groundwater 32:234–246

Yu HL, Chu HJ (2012) Recharge signal identification based on groundwater level observations. Environ Monit Assess 184:5971–5982

Yurekli K, Kurunc A, Cevik O (2004) Simulation of drought periods using stochastic models. J Eng Environ Sci 28(3):181–190