Title

The Forecasting of Groundwater Fluctuations using Time Series Analysis and Combination of Data-Driven Models

Authors

Amirhossein Najafabadipour [1]. Gholamreza Kamali [1]. Hossein Nezamabadi-pour [2]

Authors' names and Affiliations

Amirhossein Najafabadipour [1]. Gholamreza Kamali [1]. Hossein Nezamabadi-pour [2]

[1] Mining Engineering Department, Shahid Bahoonar University of Kerman, Kerman, Iran

[2] Electrical Engineering Department, Shahid Bahoonar University of Kerman, Kerman, Iran

Corresponding author

Amirhossein Najafabadipour, Mining Engineering Department, Shahid Bahoonar University of Kerman, Kerman, Iran. Email: najafabadipour@eng.uk.ac.ir

Running title

The Forecasting of Groundwater Fluctuations using Time Series Analysis

Abstract

The Forecasting of Groundwater Fluctuations is a useful tool for managing groundwater resources in the mining area. Water resources management requires identifying potential periods for groundwater drainage to prevent groundwater from entering the mine pit and imposing high costs. In this research, Auto-Regressive Integrated Moving Average (ARIMA) and Holt-Winters Exponential Smoothing (HWES) data-driven models were used for short-term modeling of the groundwater fluctuations in a piezometer around the Gohar Zamin Iron Ore Mine. For this purpose, 250 non-seasonal groundwater fluctuations data in the period 22-Nov-2018 to 29-Jul-2019, 200 data for modeling, and 50 data for prediction were used. To take advantage of all the features of the two developed models, the predictions are combined with different methods and specific weights. The results show better accuracy for the ARIMA method between the two short-term forecasts, while the HWES method requires less time for modeling. Also, among all the predictions made, the highest accuracy for the combined least-squares method is for forecasting the groundwater fluctuations in the short-term. All the forecasts show a decrease in the groundwater fluctuations, indicating pumping wells around the Gohar Zamin Iron Ore Mine area.

Keywords: Groundwater Fluctuations, ARIMA, HWES, Data-Driven Models, Combining Forecasts

1. Introduction

Groundwater is a natural resource that has costly adverse effects on mining operations (Brawner, 1986). Due to the increased depth of mining, excavation may be done below the water table, which leads to the movement of water towards mining works. Too much water entering the mining environment may delay the project or impede production, in addition to causing environmental and safety problems (Singh & Atkins, 1985). Increased equipment failure, lack of access to part of the mining area, increased use of explosives, loading problems, unsafe working conditions, and a damaging effect on pit slope stability are among the undesirable impacts of groundwater entering the mining environment. Therefore, to overcome these problems, it is necessary to develop an efficient dewatering system that predicts groundwater fluctuations that can contribute significantly to this design.

In recent years, data-driven techniques have been a useful tool for predicting groundwater fluctuations as an alternative to physical models. Lack of need for information about aquifer hydrogeological parameters is the most crucial advantage of data-driven models over physical models (Adamowski & Chan, 2011), but data-driven models need precise calibration ((Burrows & Doherty, 2015); (Xu & Valocchi, 2015); (Burrows & Doherty, 2016); (Stefania et al., 2018); (Gianni, Doherty, & Brunner, 2019); (Pan et al., 2019); (Nogueira, Schmidt, Trauth, & Fleckenstein, 2021)). Hence, data-driven models can be more appropriate if comprehensive information about subsurface characteristics is inaccessible. Also, data-driven models are superior to physical models for modeling and predicting groundwater fluctuations because they can overcome the uncertainty of parameters and data constraints ((Maskey, Dibike, Jonoski, & Solomatine, 2000); (Nikolos, Stergiadi, Papadopoulou, & Karatzas, 2008); (Banerjee, Prasad, & Singh, 2009); (Yoon, Jun, Hyun, Bae, & Lee, 2011); (Shiri, Kisi, Yoon, Lee, & Hossein Nazemi, 2013); (Boggs, Van Kirk, Johnson, & Fairley, 2014); (Chang, Wang, & Mao, 2015); (Curtis, Li, Liao, & Lusch, 2018); (Ross, Ali, Spence, Oswald, & Casson, 2019)). The extensive and successful application of data-driven models in hydrological and hydrogeological fields to predict groundwater fluctuations changes have been demonstrated in many previous studies (Guo, Song, Shi, & Li, 2020), such as runoff (Pektaş & Kerem Cigizoglu, 2013), arid region (Mirzavand & Ghazavi, 2015), semiarid (Choubin & Malekian, 2017), Spring Discharge, river basin (Gibrilla, Anornu, & Adomako, 2018).

Among data-driven models, linear models are one of the most widely used methods for predicting groundwater fluctuations. The need for less computational time and effort for training is an advantage of linear models. Struggling to handle non-linearity is the principal disadvantage of this model. To develop linear models:

I) The input time series must be stationary.

II) Residuals must be white noise.

There should be no cross-correlation and no autocorrelation between the input and the residuals; also, the mean of time series should be zero (Zanotti et al., 2019).

Developing data-driven models for accurate and efficient estimation requires a large amount of time series data, which is a time-consuming and costly process. The primary purpose of this research is to develop and compare statistical methods to forecast groundwater fluctuations using a small amount of time series data. For this purpose, Auto-Regressive Integrated Moving Average (ARIMA) and Holt-Winters Exponential Smoothing (HWES) linear statistical methods have been used to predict the fluctuations of the groundwater fluctuations of Gohar Zamin Iron Ore Mine in a short period of 250 days. The innovation of this research is the combination of developed models to cover and capture the desired features of each forecasting method.

The rest of this research is organized as follows: Section 2 describes information on the study area and developed models, Section 3 provides the results and discussion of this research, and section 4 presents the conclusions of this research.

2. Materials and Methods

2.1 Study area

Gol Gohar iron ore deposit is one of the most popular pivot points of the mining industry in the Middle East, with six separate anomalies. This deposit with a reserve of about 1200 million tons is located in an area of approximately 10 km in length and nearly 4 km in width. The mining area is generally covered with recent alluvium. Stone outcrops contain Paleozoic metamorphic rock in the south, Mesozoic and Cenozoic sedimentary rock in the east of the mine. Six separate Gol Gohar anomalies of iron deposits are situated in the metamorphic complex. The lower part of Gneiss consists of Mica schist, Amphibolite, and Quartz schist. Deposits of sediments are mostly related to Pliocene and Quaternary periods containing conglomerate, sandstone, marl, Gypsum-bearing clay, and limestone. From the tectonic point of view, Gol Gohar is located near the Zagros fault, and also, Deh Bid and Deh Shir faults located in this area are in contact with each other. The area is affected by compressive tectonics. The original faults are generally the reverse types, and the normal faults are the secondary ones. All the Quaternary faults are normal and active faults, and they can move a distance of a few centimeters to several meters. These faults play a part in guiding the water surface and groundwater through Gohar Zamin Iron Ore Mine. These young faults are created by movements of Strike-slip faults from the direction of NW-SE and have bent to the left.

In anomaly NO.3 (Gohar Zamin Iron Ore Mine), groundwater does not enter the pit; therefore, water permeates through the alluvium of the pit's stairs. Artesian explored borehole is another problem of this mine. One of the probable factors of going groundwater inflow into Gohar Zamin Iron Ore Mine is Kheyrabad plain with alluvial sediments situated in the northeast of the mine at 15 km (figure 1). Furthermore, since the mine is surrounded by various faults such as Kheyrabad, Gol Gohar, the probability of the water inflow from the Kheyrabad plain to the mine is doubled. When the mine is boring, the hydraulic gradient of groundwater, which goes through the Kheyrabad plain, changes its direction and goes through the mine. Around the Gohar Zamin Iron Ore Mine, water pumping wells are located around anomaly No. 1, which is considered a discharge area. The data set includes daily groundwater fluctuations data measured at six piezometers around the mine pit (22-Nov-2018 to 29-Jul-2019).

Fig. 1. Study area and location of Gohar Zamin Iron Ore Mine.

2.2 Time Series Analysis

A time series is a sequence of observations with an equal time interval related to one or more variables and recorded as a random sequence in time (Yoon, Hyun, Ha, Lee, & Kim, 2016); (Nixdorf & Trauth, 2018). Series data has a definite beginning and end and is not independent when the same desirable feature is used for modeling. In time series analysis, the variable changes are explained based on the current and past changes of the other variables (Palit & Popovic, 2005). The purpose of time series analysis:

1. Create input models that express the random behavior of a variable over time using a probabilistic model

2. Prediction to obtain values in forwarding time steps using the developed model

The most critical aspect of the series analysis is when it helps eliminate or reduce the inherent distribution of fluctuating components in measured values.

2.2.1 Non-Seasonal Auto-Regressive Integrated Moving Average (ARIMA)

Due to the lack of the necessary theory for the time series, statistical concepts are used based on the distribution functions to extract the essential regression to model the time series. Besides, standard regression cannot be used due to the serial correlation of time series due to the inability to estimate the coefficients of the regression equation. Box and Jenkins' approach (1976) and Auto-Regressive Integrated Moving Average (ARIMA) modeling have been used to solve this problem. The Box and Jenkins method is not a straightforward process but a series of repetitive actions, as shown in Figure 2.

Fig. 2. Box and Jenkins model construction method (Palit & Popovic, 2005).

In the model identification phase, the number of parameters of the developed mathematical model is carefully determined with the studied time series data. In the model estimation phase, the values of the model parameters are calculated by minimizing the sum of the remaining squares. In the model evaluation phase, the model's precision is examined, and the model is developed.

The essential condition for modeling and predicting a time series using the Box and Jenkins approach is the time series is stationary. The time series have a non-stationary trend, so by identifying the systematic pattern and identifying the type, and eliminating the trend, the time series is de-trending and stationarity. Locally weighted regression and smoothing scatter plot, a particular form of the nearest neighbor fit that is a non-parametric approach introduced by Cleveland (1979), can be used to identify the type of time series trend. In this method, instead of considering all the samples, a part of it is recognized. For each observation in its neighborhood, a multi-sentence regression of locally weighted fit is observed. Points with more distance weigh less, and points with less distance weigh more.

The time series converted to stationarity by differencing the time series d times. Differencing is an operation by which a replacement time series is made by taking the differences of values, like $y(t)-y(t-1)$, with the non-stationary time series pattern. During this research, the Mann-Kendall test was accustomed to assess the stationarity of the groundwater fluctuations. Therefore, the time series may be considered stationary with a significance level of 0.05, meaning that they might be used without transformation (Zanotti et al., 2019). The residuals are identified as a de-trending time series (Palit & Popovic, 2005).

In the acronym Auto-Regressive Integrated Moving Average (ARIMA), the letter I symbolize integration. Within the model identification phase, additionally to the Integrated order identification, the quantity of required AR and MA parameters for the model should be identified. The general convention for outlining the structure of ARIMA models is ARIMA $(p.d.q)$, p and q are the number of Auto-Regressive parameters and the number of moving-average parameters, respectively (Kabir et al., 2018). d stands that the number of differencing passes (Bhardwaj, Chandrasekhar, Padiyar, & Gadre, 2020). The equation of the ARIMA model for a non-seasonal time series model as:

$$y_t = c + \varnothing_1 y_{d(t-1)} + \varnothing_p y_{d(t-p)} + \ldots + \theta_1 e_{t-1} + \theta_q e_{t-q} + e_t \tag{1}$$

Where y_t is groundwater fluctuations value at time t. y_d means y differenced d times. e_t is the error of the model as a mixture of the previous error. $\varnothing_k . \theta_k$ are the Auto-Regressive and moving average constant coefficients of the model at lag k.

Due to the uncertainty of the number of parameters identified for the Auto-Regressive Integrated Moving Average (ARIMA) model, information criterion methods such as Akaike (1987)

Information Criterion (AIC), Bayesian (Gideon, 1978) Information Criterion (BIC), and Hannan and Quinn (1979) Information Criterion (HQIC) can be used to estimate the optimal number of parameters. After determining the required number of parameters of the model, the model should be evaluated using these values. Some unique statistical methods, such as Maximum Likelihood, can be used to determine the model. The assessed demonstrate at this step ought to have the most reduced mistake rate than the real proof.

The validity of the developed model must be confirmed in the last phase of the model construction. The fit of the developed model should be checked with time series data. In addition, future forecast values should be close to the actual values. A simplified approach to verify the minimum number of model parameters needed to represent the observation data is to check the residuals' mutual noncorrelation. In the case of residual correlation, the number of model parameters must be increased. The residual diagnostic methods are appropriate to verify this correlation. Residual diagnostics includes statistical calculations of the autocorrelation of residuals. The correlogram of residuals is assessed for checking the mutual correlation of residuals. Spikes in the correlogram signs that the residuals may be correlated, which the model created is not satisfactory.

2.2.2 Holt-Winters Exponential Smoothing (HWES)

The Exponential Smoothing approach is especially suitable for short-term forecasting. In this method, weight factors are used for the past values, and the weight factors are reduced in the form of a view with a distance from the previous values of the time series (Bowerman & O'Connell, 1993). This approach makes possible formula of the prediction algorithm that requires only a few recent data with fewer calculations (Holt, 2004).

The smoothed series y_t is given by:

$$y_{t+k}=a+bk \tag{2}$$

a is the permanent component, and b is the trend. The following recursions characterize these two coefficients:

$$a(t)=\alpha\, y_t+(1-\alpha)\left(a(t-1)+b(t-1)\right) \tag{3}$$

$$b(t)=\beta\left(a(t)-a(t-1)\right)+1-\beta b(t-1) \tag{4}$$

Where α and β are damping factors fluctuating between 0 and 1 depending on the specific time series characteristics. Time series patterns and smoothing objectives can affect the optimal value of these parameters. α values between 0.1 and 0.3 are more commonly used because they rely on many last observations to predict. Values close to 1 are seldom used because they make predictions that depend on recent data. For example, $\alpha=1$ shows the forecast is equal to recent observations.

2.2.3 Forecast Combining

There are several ways to predict the time series, but choosing the best method is difficult. Besides, each of the available techniques offers different predictive results, which after comparing the results, the best model is selected based on the professional experience of the user. These models and forecasting, in addition to different fundamental assumptions, also use additional information. Traditionally, an individual prediction was chosen with the least error. Various studies have shown that forecast combining has a more precise result than selecting the best forecast. Timmermann (2006) presents an overview of such studies, citing Clemen (1989) and Makridakis and Hibon (2000) specifically.

Forecast combining can be combining multiple forecasts into one forecast. The forecast combining method is based on forecast evaluation and model selection techniques. To choose the most practical forecasting, the acceptability of each forecasting must be measured. Instead of identifying the best forecasting, the feature of all predictions can be used by generating weighting coefficients. Combined forecasting methods regularly utilize a simple weighted average calculation. The cross-forecast average with various predictions assigned different weights is computed. Different weights are used to determine complex weight schemes during the forecast period. The following is a list of commonly used methods that use constant weights through time:

Simple mean: In this method, the arithmetic mean of forecasts at each observation within the forecast sample is calculated. During this method, every forecast is given an identical weight.

Simple median: The median of the prediction at each observation within the prediction sample will be estimated for the simple median method. The implicit (0, 1) weights are time-varying as every forecasting method is also the median for a few observations.

Least-squares weights: To use this method, the actual values of the forecasted variable for a few of the forecast periods must be identified. This method is calculated by regressing the forecasts against particular values and using regression coefficients as weights. In the least-squares method, the underlying individual forecasts can be unbiased, and also the resulting average can be outside the scope of the underlying forecasts.

Mean squares error weights: Mean squares error (MSE) weighting during the forecast period compares the actual values with the individual forecasts (Stock & Watson, 2002). After computing each forecast, individual forecast weights are formed using the following equation.

$$w_i = \frac{\frac{1}{MSE_i^k}}{\sum_{j=1}^{N} \frac{1}{MSE_j^k}} \tag{5}$$

k is employed to change the MSE to different powers. The most comprehensive power for k is 1.

MSE ranks: This method estimates the mean squares error of every forecast, then ranks them and finally calculates the ratio of the inverse of the ranks (Aiolfi, Capistran, & Timmermann, 2010). The weight of forecasting is obtained by dividing its rank by the sum of all forecast ranks.

Smoothed AIC weights: This method applies the Akaike (1987) information criterion from the developed model that created each forecast. The weight of forecasting calculates from the following equation:

$$w_i = \frac{\exp\left(-0.5\, AIC_i\right)}{\sum_{j=1}^{N} \exp\left(-0.5\, AIC_j\right)} \tag{6}$$

Approximate Bayesian model averaging weights: Bayesian model averaging (BMA) weights use the Bayesian information criterion from the developed model that produced every forecast or the first estimation, then the weight of forecasting calculates from the following equation:

$$w_i = \frac{\exp\left(-0.5\, BIC_i\right)}{\sum_{j=1}^{N} \exp\left(-0.5\, BIC_j\right)} \tag{7}$$

2.2.4 Forecast Evaluation Statistics

To describe the error related to the model output, various statistical measures can be used to compare the effectiveness of the developed models. The performance of the trained model is compared in terms of statistical measurement of precision. During this research, the Root Mean Squared Error (RMSE), the Mean Absolute Error (MAE), the Mean Absolute Percentage Error (MAPE), and Theil Inequality Coefficient are taken under consideration to check the efficiency of the models as predictive tools. RMSE and MAE are a measure of the quality deviation of the residuals.

Supposes, the forecast sample is $j=T+1.T+2.\ldots.T+h$, and in period t, y_t and $\hat{y}_t$ denote the particular and forecasted groundwater fluctuations, respectively. Forecast evaluation is calculated from the following equations:

$$RMSE=\sqrt{\sum_{t=T+1}^{T+h}\frac{(\hat{y}_t-y_t)^2}{h}} \tag{8}$$

$$MAE=\sum_{t=T+1}^{T+h}\frac{|\hat{y}_t-y_t|}{h} \tag{9}$$

$$MAPE=100\sum_{t=T+1}^{T+h}\frac{\left|\frac{\hat{y}_t-y_t}{y_t}\right|}{h} \tag{10}$$

$$Theil\,Inequality\,Coefficient=\frac{\sqrt{\sum_{t=T+1}^{T+h}\frac{(\hat{y}_t-y_t)^2}{h}}}{\sqrt{\sum_{t=T+1}^{T+h}\frac{\hat{y}_t^{\,2}}{h}}+\sqrt{\sum_{t=T+1}^{T+h}\frac{y_t^{\,2}}{h}}} \tag{11}$$

h is the total amount of data within the sequence. The developed model performs better with smaller RMSE and MAE, MAPE, and Theil Inequality Coefficient values (Zanotti et al., 2019). To evaluate the models' performance, different performance criteria can be used, as different criteria, with emphasis on different aspects of the model's predictive power, examine the performance of the model (Maier, Jain, Dandy, & Sudheer, 2010). In this research, the Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan and Quinn Information Criterion (HQIC) are calculated:

$$AIC = nlog\left(\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{n}\right) + 2p \tag{12}$$

$$BIC = nlog\left(\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{n}\right) + p\ln\left(n\right) \tag{13}$$

$$HQIC = nlog\left(\frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{n}\right) + 2pln\left(\ln\left(n\right)\right) \tag{14}$$

y_i is the i-th target value of groundwater fluctuations, $\hat{y}_i$ is that the i-th groundwater fluctuations value estimated by the developed model, and n is that the total amount of data within the sequence. The scalar p is the number of parameters of the developed model that are estimated during the training. Regarding the existence of two suitable models with acceptable residues, one of them, which has a lower information criterion, is selected because it is better. In addition to measuring model error, information criteria also consider model complexity. Therefore, they need the potential to end in more parsimonious models (Maier et al., 2010).

To settle on the most straightforward model from the prevailing predictions, it means which model has better predictor precision. The simple approach is to choose the forecast that has the smaller error measurement supported one in every of the error measurements described in forecasting errors. The Diebold and Mariano (2002) test is used to check the precision of several competing predictions. For one-step-ahead prediction, the test statistic is computed as:

$$S=\frac{\overline{d}}{s_d} \tag{15}$$

Where

$$d=L_1-L_2 \tag{16}$$

And L_i, $i=1.2$ is either an absolute or squared difference between the forecast and also the actual,

$$L_i=\left|\hat{y}_i-y\right| \tag{17}$$

Or

$$L_i=\left(\hat{y}_i-y\right)^2 \tag{18}$$

where $\overline{d}$ and s_d are the mean and sample standard deviation of d.

A combination test or Forecast Encompassing Test is used to test whether the average or combination of predictions performed is better than the individual predictions (Chong and Hendry (1986) and refined by Timmermann (2006)). This test combines several forecasting into single forecasting. In this test, if a forecast includes all information contained within the other individual forecasts, that forecast will be even as good as a combination of all of the forecasts. By performing regression of the model, a test of this hypothesis is performed:

$$Y_{t+h}-\hat{Y}_{t+h.i}=\beta_0+\sum_{j\neq i}^{N}\beta_j\hat{Y}_{t+h.j} \tag{19}$$

For forecast *i*, during the forecast period, Y_{t+h} is the vector of actual values, and during the identical period, $\hat{Y}_{t+h.i}$ is the vector of forecast values. A test for whether forecast *i* contains all the information of the other forecasts is also performed by testing whether $\beta_j = 0. \forall (j \neq i)$; if the difference between the actuality values and, therefore, the forecasted values from forecast *i* are not associated with the forecasts from all other developed models, then forecast *i* are often used individually. If the differences are laid low with the different forecasts, then the recent forecasts should be contained within the formation of a composite forecast.

In this research, to predict the time series using the Box and Jenkins approach, the groundwater fluctuations data was first de-trended using the method introduced by Cleveland (1979). Then the optimal model for Auto-Regressive Integrated Moving Average (ARIMA) modeling is identified, and the groundwater fluctuations for the future are predicted using it. Using the Holt-Winters Exponential Smoothing model, the groundwater fluctuations are forecasted, and the error of different methods is calculated. To benefit from the desirable features of each of these predictions, the candidate models are combined, and the combined models are evaluated. All models and evaluations in the Eviews software package are provided.

3. Results and Discussion

3.1 Auto-Regressive Integrated Moving Average (ARIMA) model

To stationarity the time series, using the Box and Jenkins approach and the Auto-Regressive Integrated Moving Average (ARIMA) model, time series trends must be identified and eliminated. To study groundwater fluctuations trend around the study area of the available six piezometer data, piezometer records No. 2 located in the western part of Pit Mining from 22-Nov-2018 to 29-Jul-2019 were examined. The statistical explanation of the groundwater fluctuations is shown in Table 1.

Table 1

The statistical explanation of the groundwater fluctuations

In visual inspection, the line fitted to groundwater data using the nearest neighbor fit method indicates that the trend of the time series is decreasing and non-seasonal. The declining time series shows the groundwater withdrawal by pumping wells around the mine pit (Figure 3). In the first 90 days, the average groundwater fluctuations are about 1613.85 meters above sea level because the pumping wells around the mine pit have been more active. During 90 to 160 days, the average groundwater fluctuations are about 1611.95 meters above sea level, which indicates less activity of pumping wells. In the last 90 days, again, with more pumping activity, the average groundwater fluctuations are about 1611.06 meters above sea level.

Fig. 3. The fitted line to groundwater fluctuations.

After identifying the trend to remove it, the logarithm and the first difference is taken from the time series. A value close to 1 for the p-value in the Mann-Kendall test indicates eliminating trends in groundwater fluctuations time series data. After removing the trend, the information criterion was used to identify the orders in the Box and Jenkins approach. Figure 4 shows the results of the top 20 Auto-Regressive Integrated Moving Average (ARIMA) models for different orders, using Akaike information criteria, Bayesian information criteria, and Hannan and Quinn information criteria. For Akaike Information Criterion (AIC), Bayesian Information Criterion (BIC), and Hannan and Quinn Information Criterion (HQIC) ARIMA models (6,1,3), (1,1,1), and (1,1,1) are suitable candidates for modeling, respectively. According to the simple parsimony, the ARIMA model (1,1,1) has been used for modeling and prediction.

Fig. 4. AIC, BIC, and HQIC values for the top 20 models to estimate the model orders.

In the modeling phase using the Box and Jenkins approach, using the candidate model, out of 250 daily data of groundwater fluctuations available, 200 samples were modeled using the maximum likelihood method. The fitted model and the residual value for the estimated model are shown in Figure 5. Given that most of the remaining value is in the confidence range, as well as the average close to zero (-6.05E-13), it can be concluded that for the developed model, the pattern is well identified, and a passable estimate has been made for it.

Fig. 5. Actual, fitted, and residual graphs for the first difference ARIMA model.

In the evaluation phase of the estimated model using the Box and Jenkins approach, partial autocorrelation function (PACF) and autocorrelation function (ACF) of the theoretical residues (red) and the estimated model (blue) compares for 24 lags (Figure 6). According to the LJUNG and BOX (1978) statistics, the model does not have a linear dependence. The serial correlation is well captured, which indicates an acceptable estimate for the estimated model.

Fig. 6. ACF and PACF of the theoretical residues and the estimated model.

The next step is to static forecast the groundwater fluctuations for the next 50 days using the estimated Auto-Regressive Integrated Moving Average (ARIMA) model. The prediction accuracy of the time series is shown in Table 1, which indicates a low error. The prediction made by the ARIMA method indicates a decrease in the groundwater fluctuations in the next 50 days. Groundwater fluctuations decreased sharply between 6 Jul-2019 and 13-Jul-2019, indicating more pumping activity. In the period from 14-Jul-2019 to 29-Jul-2019, the slope of the groundwater reduction chart has reduced. On the last day of the forecast, the groundwater fluctuations are estimated at 1607.64 meters, which differs from the actual value of 1607.46 by about 0.18 meters.

3.2 Exponential Smoothing Model

In order to predict the time series of groundwater fluctuations, the Holt-Winters parameters for $\alpha = 0.96$, $\beta = 0.33$ and $\gamma = 0$ are estimated. α Nearly one indicates that groundwater fluctuations forecasting is commonly based on recent values in the time series. A value of 0.33 for β demonstrates linearity, and a value of 0 for γ shows that the time series is non-seasonal. The average amount of groundwater fluctuations and the trend for modeling are 1611.57 and -0.04, respectively. Figure 7 shows the estimated model using the Holt-Winters Exponential Smoothing (HWES) method with its residuals. Most of the estimated residuals of the time series are in the confidence range, and close to zero (-9.83E-13) indicates that the model is acceptable.

Fig. 7. Actual, fitted, and residual graphs for Holt-Winters Exponential Smoothing model.

In the predicted model using Holt-Winters Exponential Smoothing (HWES), the groundwater fluctuations have decreased, and the accuracy of the forecast for this model is shown in Table 2. The lower the calculated error value, the better the ability to predict the developed model according to that criterion. The values of these statistics were 0.107297 and 0.069369 for RMSE(m) and MAE(m), respectively, indicating there was some minor underestimation of the observed groundwater fluctuations. Theil inequality coefficient and the Mean Absolute Percent Error (MAPE) are scale invariants. The Theil inequality coefficient lies between 0 and 1, where 0 indicates a perfect fit. A low value of MAPE (here 0.004311 %) and the Theil inequality coefficient (here 3.33E-05) is considered a good indicator. In the forecast made by the Holt-Winters Exponential Smoothing (HWES) method, the groundwater fluctuations on 29-Jul-2019 are estimated at 1607.62 meters, which differs from the real value of 1607.46 by about 0.16 meters.

Table 2

Forecast evaluation for nine forecasts applied on groundwater fluctuations data in Gohar Zamin Iron Ore Mine

3.3 Combined Forecast

To evaluate the accuracy of the two predictions Holt-Winters Exponential Smoothing (HWES) and Auto-Regressive Integrated Moving Average (ARIMA), the Diebold and Mariano test was used, which for absolute and squared error, the p-value is 0.21 and 0.72, respectively. The p-value above 5% of the null hypothesis is accepted and indicates that the two predictions have the same accuracy. Since the forecasts made by ARIMA and HWES models have the same efficiency and each has its desired characteristics, to capture and cover all the forecasting features, combining two forecasts has been used in different methods. The Encompassing test, introduced by Chong and Hendry (1986) and refined by Timmermann (2006), was used to test whether the prediction combination performed better than individual forecasts. The p-value of about 0.66 for ARIMA prediction indicates a higher accuracy of this method than HWES.

To combine the predictions, 250 available data, 200 data for training, and 50 data for evaluation were used. The evaluation results of the composition of the predictions made by different methods are shown in Table 2. The results show the highest accuracy related to the least-squares combination method. Table 3 calculates the weights obtained to combine the Holt-Winters Exponential Smoothing (HWES) and Auto-Regressive Integrated Moving Average (ARIMA) predictions.

Table 3

The weights obtained to combine the predictions of the ARIMA and HWES models

Figure 8 shows a forecast comparison graph for the actual mounting of groundwater fluctuations, Holt-Winters Exponential Smoothing (HWES), and Auto-Regressive Integrated Moving Average (ARIMA) forecasts, as well as different methods of combining forecasts. The predictions show the high accuracy of the ARIMA method compared to the HWES method and the least-squares method in combining models.

Fig. 8 Time series forecast by actual groundwater fluctuations, ARIMA, HWES, and combined forecast.

4. conclusion

Groundwater fluctuations for mines can be forecasted by short-term analysis that shows the impact of groundwater recharging and harvesting from water storage in an aquifer. The records of piezometer No. 2 in the western part of Gohar Zamin Iron Ore Mine were reviewed from 22-Nov-2018 to 29-Jul-2019 to study groundwater fluctuations. In this research, the predictive ability of Holt-Winters Exponential Smoothing (HWES) and non-seasonal Auto-Regressive Integrated Moving Average (ARIMA) models at the time series of daily groundwater fluctuations around the Gohar Zamin Iron Ore Mine pit was compared.

Since each prediction has its characteristics, this research showed that combining predictions made with different methods has better results than individual predictions. Diebold and Mariano’s test with a p-value above 5% indicates a similar accuracy of the two predictive methods. An encompassing test with a p-value of about 0.66 for Auto-Regressive Integrated Moving Average (ARIMA) prediction indicates a higher efficiency of this method than Holt-Winters Exponential Smoothing (HWES) in short-term predictions. Also, the results of the prediction combination show a higher accuracy of the least-squares combination method than other prediction methods. Forecasts made by simulation of time series indicated that if the current situation is still around the pit, it is predicted that the groundwater fluctuations will decrease, indicating the activity of pumping wells around the pit. Due to the acceptable results of the prediction combination, which shows the high accuracy of this method, it is recommended to predict the short-term time series.

Acknowledgments

The authors would like to acknowledge the cooperation of the Gohar Zamin Iron Ore Mine.

References

Adamowski, J., & Chan, H. F. (2011). A wavelet neural network conjunction model for groundwater level forecasting. *Journal of Hydrology, 407*(1), 28-40. doi:https://doi.org/10.1016/j.jhydrol.2011.06.013

Aiolfi, M., Capistran, C., & Timmermann, A. (2010). Forecast combinations. *CREATES research paper*(2010-21).

Akaike, H. (1987). Factor analysis and AIC. In *Selected papers of hirotugu akaike* (pp. 371-386): Springer.

Banerjee, P., Prasad, R. K., & Singh, V. S. (2009). Forecasting of groundwater level in hard rock region using artificial neural network. *Environmental Geology, 58*(6), 1239-1246. doi:10.1007/s00254-008-1619-z

Bhardwaj, S., Chandrasekhar, E., Padiyar, P., & Gadre, V. M. (2020). A comparative study of wavelet-based ANN and classical techniques for geophysical time-series forecasting. *Computers & Geosciences, 138*, 104461. doi:https://doi.org/10.1016/j.cageo.2020.104461

Boggs, K. G., Van Kirk, R., Johnson, G. S., & Fairley, J. P. (2014). Forecast of Natural Aquifer Discharge Using a Data-Driven, Statistical Approach. *Groundwater, 52*(6), 853-863. doi:https://doi.org/10.1111/gwat.12133

Bowerman, B. L., & O'Connell, R. T. (1993). *Forecasting and Time Series: An Applied Approach*: Duxbury Press.

Box, G. E. P., Jenkins, G. M., & Day, H. (1976). *Time Series Analysis: Forecasting and Control*: Holden-Day.

Brawner, C. O. (1986). Groundwater and coal mining. *Mining Science and Technology, 3*(3), 187-198. doi:https://doi.org/10.1016/S0167-9031(86)90331-2

Burrows, W., & Doherty, J. (2015). Efficient Calibration/Uncertainty Analysis Using Paired Complex/Surrogate Models. *Groundwater, 53*(4), 531-541. doi:https://doi.org/10.1111/gwat.12257

Burrows, W., & Doherty, J. (2016). Gradient-based model calibration with proxy-model assistance. *Journal of Hydrology, 533*, 114-127. doi:https://doi.org/10.1016/j.jhydrol.2015.11.033

Chang, J., Wang, G., & Mao, T. (2015). Simulation and prediction of suprapermafrost groundwater level variation in response to climate change using a neural network model. *Journal of Hydrology, 529*, 1211-1220. doi:https://doi.org/10.1016/j.jhydrol.2015.09.038

Chong, Y. Y., & Hendry, D. F. (1986). Econometric Evaluation of Linear Macro-Economic Models. *The Review of Economic Studies, 53*(4), 671-690. doi:10.2307/2297611

Choubin, B., & Malekian, A. (2017). Combined gamma and M-test-based ANN and ARIMA models for groundwater fluctuation forecasting in semiarid regions. *Environmental Earth Sciences, 76*(15), 538. doi:10.1007/s12665-017-6870-8

Clemen, R. T. (1989). Combining forecasts: A review and annotated bibliography. *International Journal of Forecasting, 5*(4), 559-583. doi:https://doi.org/10.1016/0169-2070(89)90012-5

Cleveland, W. S. (1979). Robust locally weighted regression and smoothing scatterplots. *Journal of the American statistical association, 74*(368), 829-836.

Curtis, Z. K., Li, S.-G., Liao, H.-S., & Lusch, D. (2018). Data-Driven Approach for Analyzing Hydrogeology and Groundwater Quality Across Multiple Scales. *Groundwater, 56*(3), 377-398. doi:https://doi.org/10.1111/gwat.12584

Diebold, F. X., & Mariano, R. S. (2002). Comparing predictive accuracy. *Journal of Business & economic statistics, 20*(1), 134-144.

Gianni, G., Doherty, J., & Brunner, P. (2019). Conceptualization and Calibration of Anisotropic Alluvial Systems: Pitfalls and Biases. *Groundwater, 57*(3), 409-419. doi:https://doi.org/10.1111/gwat.12802

Gibrilla, A., Anornu, G., & Adomako, D. (2018). Trend analysis and ARIMA modelling of recent groundwater levels in the White Volta River basin of Ghana. *Groundwater for Sustainable Development, 6*, 150-163. doi:https://doi.org/10.1016/j.gsd.2017.12.006

Gideon, S. (1978). Estimating the Dimension of a Model. *The Annals of Statistics, 6*(2), 461-464. doi:10.1214/aos/1176344136

Guo, T., Song, S., Shi, J., & Li, J. (2020). Groundwater Depth Forecasting Using Configurational Entropy Spectral Analyses with the Optimal Input. *Groundwater, 58*(5), 749-758. doi:https://doi.org/10.1111/gwat.12968

Hannan, E. J., & Quinn, B. G. (1979). The Determination of the Order of an Autoregression. *Journal of the Royal Statistical Society: Series B (Methodological), 41*(2), 190-195. doi:https://doi.org/10.1111/j.2517-6161.1979.tb01072.x

Holt, C. C. (2004). Forecasting seasonals and trends by exponentially weighted moving averages. *International Journal of Forecasting, 20*(1), 5-10. doi:https://doi.org/10.1016/j.ijforecast.2003.09.015

Kabir, F., Yu, N., Yao, W., Wu, L., Jiang, J. H., Gu, Y., & Su, H. (2018). Impact of aerosols on reservoir inflow: A case study for Big Creek Hydroelectric System in California. *Hydrological Processes, 32*(22), 3365-3390. doi:https://doi.org/10.1002/hyp.13265

LJUNG, G. M., & BOX, G. E. P. (1978). On a measure of lack of fit in time series models. *Biometrika, 65*(2), 297-303. doi:10.1093/biomet/65.2.297

Maier, H. R., Jain, A., Dandy, G. C., & Sudheer, K. P. (2010). Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions. *Environmental Modelling & Software, 25*(8), 891-909. doi:https://doi.org/10.1016/j.envsoft.2010.02.003

Makridakis, S., & Hibon, M. (2000). The M3-Competition: results, conclusions and implications. *International Journal of Forecasting, 16*(4), 451-476. doi:https://doi.org/10.1016/S0169-2070(00)00057-1

Maskey, S., Dibike, Y. B., Jonoski, A., & Solomatine, D. (2000). Groundwater model approximation with artificial neural network for selecting optimum pumping strategy for plume removal. *International Institute for Infrastructural, Hydraulic and Environmental Engineering, Delft, The Netherlands*.

Mirzavand, M., & Ghazavi, R. (2015). A Stochastic Modelling Technique for Groundwater Level Forecasting in an Arid Environment Using Time Series Methods. *Water Resources Management, 29*(4), 1315-1328. doi:10.1007/s11269-014-0875-9

Nikolos, I. K., Stergiadi, M., Papadopoulou, M. P., & Karatzas, G. P. (2008). Artificial neural networks as an alternative approach to groundwater numerical modelling and environmental design. *Hydrological Processes: An International Journal, 22*(17), 3337-3348.

Nixdorf, E., & Trauth, N. (2018). Evaluating the reliability of time series analysis to estimate variable riparian travel times by numerical groundwater modelling. *Hydrological Processes, 32*(3), 408-420. doi:https://doi.org/10.1002/hyp.11428

Nogueira, G. E. H., Schmidt, C., Trauth, N., & Fleckenstein, J. H. (2021). Seasonal and short-term controls of riparian oxygen dynamics and the implications for redox processes. *Hydrological Processes, 35*(2), e14055. doi:https://doi.org/10.1002/hyp.14055

Palit, A. K., & Popovic, D. (2005). Computational Intelligence in Time Series Forecasting: Theory and Engineering Applications.

Pan, J., Shangguan, W., Li, L., Yuan, H., Zhang, S., Lu, X., . . . Dai, Y. (2019). Using data-driven methods to explore the predictability of surface soil moisture with FLUXNET site data. *Hydrological Processes, 33*(23), 2978-2996. doi:https://doi.org/10.1002/hyp.13540

Pektaş, A. O., & Kerem Cigizoglu, H. (2013). ANN hybrid model versus ARIMA and ARIMAX models of runoff coefficient. *Journal of Hydrology, 500*, 21-36. doi:https://doi.org/10.1016/j.jhydrol.2013.07.020

Ross, C. A., Ali, G., Spence, C., Oswald, C., & Casson, N. (2019). Comparison of event-specific rainfall–runoff responses and their controls in contrasting geographic areas. *Hydrological Processes, 33*(14), 1961-1979. doi:https://doi.org/10.1002/hyp.13460

Shiri, J., Kisi, O., Yoon, H., Lee, K.-K., & Hossein Nazemi, A. (2013). Predicting groundwater level fluctuations with meteorological effect implications—A comparative study among soft computing techniques. *Computers & Geosciences, 56*, 32-44. doi:https://doi.org/10.1016/j.cageo.2013.01.007

Singh, R. N., & Atkins, A. S. (1985). Analytical techniques for the estimation of mine water inflow. *International Journal of Mining Engineering, 3*(1), 65-77. doi:10.1007/BF00881342

Stefania, G. A., Rotiroti, M., Fumagalli, L., Simonetto, F., Capodaglio, P., Zanotti, C., & Bonomi, T. (2018). Modeling groundwater/surface-water interactions in an Alpine valley (the Aosta Plain, NW Italy): the effect of groundwater abstraction on surface-water resources. *Hydrogeology Journal, 26*(1), 147-162. doi:10.1007/s10040-017-1633-x

Stock, J. H., & Watson, M. W. (2002). Forecasting Using Principal Components from a Large Number of Predictors. *Journal of the American statistical association, 97*(460), 1167-1179.

Timmermann, A. (2006). Chapter 4 Forecast Combinations. In G. Elliott, C. W. J. Granger, & A. Timmermann (Eds.), *Handbook of Economic Forecasting* (Vol. 1, pp. 135-196): Elsevier.

Xu, T., & Valocchi, A. J. (2015). Data-driven methods to improve baseflow prediction of a regional groundwater model. *Computers & Geosciences, 85*, 124-136. doi:https://doi.org/10.1016/j.cageo.2015.05.016

Yoon, H., Hyun, Y., Ha, K., Lee, K.-K., & Kim, G.-B. (2016). A method to improve the stability and accuracy of ANN- and SVM-based time series models for long-term groundwater level predictions. *Computers & Geosciences, 90*, 144-155. doi:https://doi.org/10.1016/j.cageo.2016.03.002

Yoon, H., Jun, S.-C., Hyun, Y., Bae, G.-O., & Lee, K.-K. (2011). A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. *Journal of Hydrology, 396*(1), 128-138. doi:https://doi.org/10.1016/j.jhydrol.2010.11.002

Zanotti, C., Rotiroti, M., Sterlacchini, S., Cappellini, G., Fumagalli, L., Stefania, G. A., . . . Bonomi, T. (2019). Choosing between linear and nonlinear models and avoiding overfitting for short and long term groundwater level forecasting in a linear system. *Journal of Hydrology, 578*, 124015. doi:https://doi.org/10.1016/j.jhydrol.2019.124015