Imputing missing data using grey system theory and the biplot method to forecast groundwater levels and yields

Jelena Ratković, Dušan Polomčić, Zoran Gligorić and Dragoljub Bajić*

Faculty of Mining and Geology, University of Belgrade, Đušina 7, 11000 Belgrade, Serbia;
(jmocevic45@gmail.com, dusan.polomcic@rgf.bg.ac.rs, zoran.gligoric@rgf.bg.ac.rs, *correspondence: dragoljub.bajic@rgf.bg.ac.rs)

Crossref
Similarity Check
Powered by iThenticate

Abstract

Groundwater management is one of today's important tasks. It has become necessary to seek out increasingly reliable methods to conserve groundwater resources. Dependable forecasting of the amounts of groundwater that can be abstracted in a sustainable manner requires long-term monitoring of the groundwater regime (rate of abstraction and groundwater levels). Monitoring of the groundwater source for the town of Bečej, Serbia had been disrupted for multiple years. The objective of the paper is to assess the possibility of reinterpreting the missing data or, in other words, to reconstruct the operation of the groundwater source and its effect on groundwater levels. At the Bečej source, groundwater is withdrawn from three water-bearing strata comprised of fine- to coarse-grained sands. Historic data are used to reconstruct the operation of the Bečej source between 1st of October 1980 to 1st of May 2010. The monitored parameters are total source yield and piezometric head at seven observation wells and 14 pumping wells. A data reconstruction methodology was developed, which included the use of an autoregressive (AR) model, a grey model (GM), and the biplot method. The methodology is applied to fill the data gaps during the considered period. The paper also describes the criteria for evaluating the accuracy of the AR model, GM, and biplot method. The proposed data reconstruction approach yielded satisfactory results and the methodology is deemed useful for the Bečej source data, as well as other historic data not necessarily associated with groundwater sources, but also groundwater control and protection systems, as well as hydrometeorological, hydrological and similar uses.

1. INTRODUCTION

Groundwater is one of the primary sources of water for various uses. The demand for groundwater will steadily rise in the future. This is a growing problem and more and more attention will be devoted to seeking out methods and strategies for progressively efficient groundwater management. In practice, the water supply for a particualr area is often provided by developing groundwater sources that deliver sufficient water quantities of a satisfactory quality.

Effective groundwater source management largely depends on local natural conditions, the amount and quality of collected operational data, and the efficient and reliable forecasting of the future water demand and the ability of the hydrogeological setting to meet that demand. Roughly 75% of the water demand in Serbia relies on groundwater. However, there has been no regular monitoring of surface water and groundwater dynamics during a period of economic crisis in the country, which resulted in an underdeveloped observation network. At groundwater sources, regular monitoring was hindered by poor organization and lack of instrumentation and human resources. More recently, the surface water and groundwater observation network has been expanded based on European Union recommendations and the scope and extent of monitoring are now satisfactory.

The need to develop methods that will forecast sustainable rates of groundwater abstraction is becoming increasingly pronounced. Sufficiently reliable prognostic methods would create conditions for effective conservation of groundwater resources, addressing problems such as inefficient use and over-exploitation,

which impairs groundwater quality. Reliable forecasts of groundwater levels, groundwater balance, and the velocity and direction of groundwater flow play a major role, in view of the fact that they provide important information about the characteristics of the tapped aquifer. With the exception of groundwater quality, forecasting basically focuses on groundwater levels at a particular rate of abstraction.

Today, many methods are used to *forecast groundwater levels*, such as Nonlinear Autoregressive Networks (NARX) (GUZMAN et al., 2017; WUNSCH et al., 2018), Artificial Neutral Networks (ANN) (NOURANI et al., 2008; RAKHSHANDEHROO et al., 2012; CHITSAZAN et al., 2015; MOHANTY et al., 2015), Time Series Methods (MIRZAVAND & GHAZAVI, 2015), ARIMA and RBF networks (YAN & MA, 2016), Adaptive Neuro-Fuzzy Inference System (ANFIS) (GULDAL & TONGAL, 2009; EMAMGHOLIZADEH et al., 2014), fuzzy neural networks (ALVISI & FRANCHINI, 2011), Fuzzy C-Mean Clustering and Singular Spectrum Analysis (POLOMČIĆ et al., 2017), and hydrodynamic modeling (POLOMČIĆ et al., 2011; POLOMČIĆ et al., 2014).

The objective of the paper is to present a case study of the Bečej groundwater source for drinking water supply (Serbia) and demonstrate the reconstruction of missing data (over a 10-year period), using the autoregressive model, the grey system theory and biplot method, in order to improve decision-making related to groundwater management.

Autoregressive modeling was used as a supplementary method to extract the points in time (dates) of no monitoring. It has been implemented by many other researchers in the fields of

hydrology (CHAKRABORTY et al, 2009; MEHDIZADEH, 2020), seismology (XIAOJUN et al., 2015), exodynamics (YONGPING et al., 2020), and hydrogeology (RISTIĆ VAKA-NJAC et al., 2018).

Many researchers in hydrogeology and similar disciplines have applied the *grey system theory* to address diverse problems, such as predicting earthquakes (DAOGONG et al., 1999), forecasting of pore pressure changes (MOHAMMED et al., 2010), projecting land subsidence to prevent hazards (TANG et al., 2008), sustainable groundwater management for drinking water supply (JIANFEI et al., 2012), estimating the amount of rainfall and runoff to prevent flooding (YU et al., 2001; KANG et al., 2006), forecasting groundwater levels (MAHMOD et al., 2013; KANG & MAENG, 2016) and predicting and analyzing spring discharges (HAO et al., 2010).

The *biplot method* is also applied by many researchers in different fields. For example, this statistical method has been very useful in addressing environmental protection issues, to predict the evolution of air pollution (GONZALEZ CABRERA et al., 2006; ZHANG et al., 2018); in mechanics, to solve problems applying SVD (singular value decomposition, which is the basis of

the biplot method) (MEIJAARD, 1993; GAI & HU, 2018); in geophysics, to improve the quality of seismic data (MARTINS et al., 2016) and to determine the spatial relationship between deep geologic structures and gold (ZHAO & CHEN, 2010).

Upon reconstruction of the missing piezometric head data and total groundwater source yields, the gaps in the Bečej groundwater source (Bečej GS) data were filled over the study period. Sufficiently reliable input data were provided to assess the feasibility of Bečej GS expansion, predict the long-term effect of groundwater abstraction, and propose optimal solutions for future production wells.

2. STUDY AREA

The study area is the extended area of Bečej GS, Serbia, as shown in Fig. 1. The groundwater source is located west of the city of Bečej. Groundwater from the tapped aquifers is used for both the municipal and industrial water supplies. Between 1980 and 2010, 25 production wells and seven observation wells were drilled at Bečej GS. Fourteen production wells are currently in service (Fig. 1). The others have been decommissioned. The wells tap several aquifers. Their individual discharge capacities are 3 – 16 l/s.
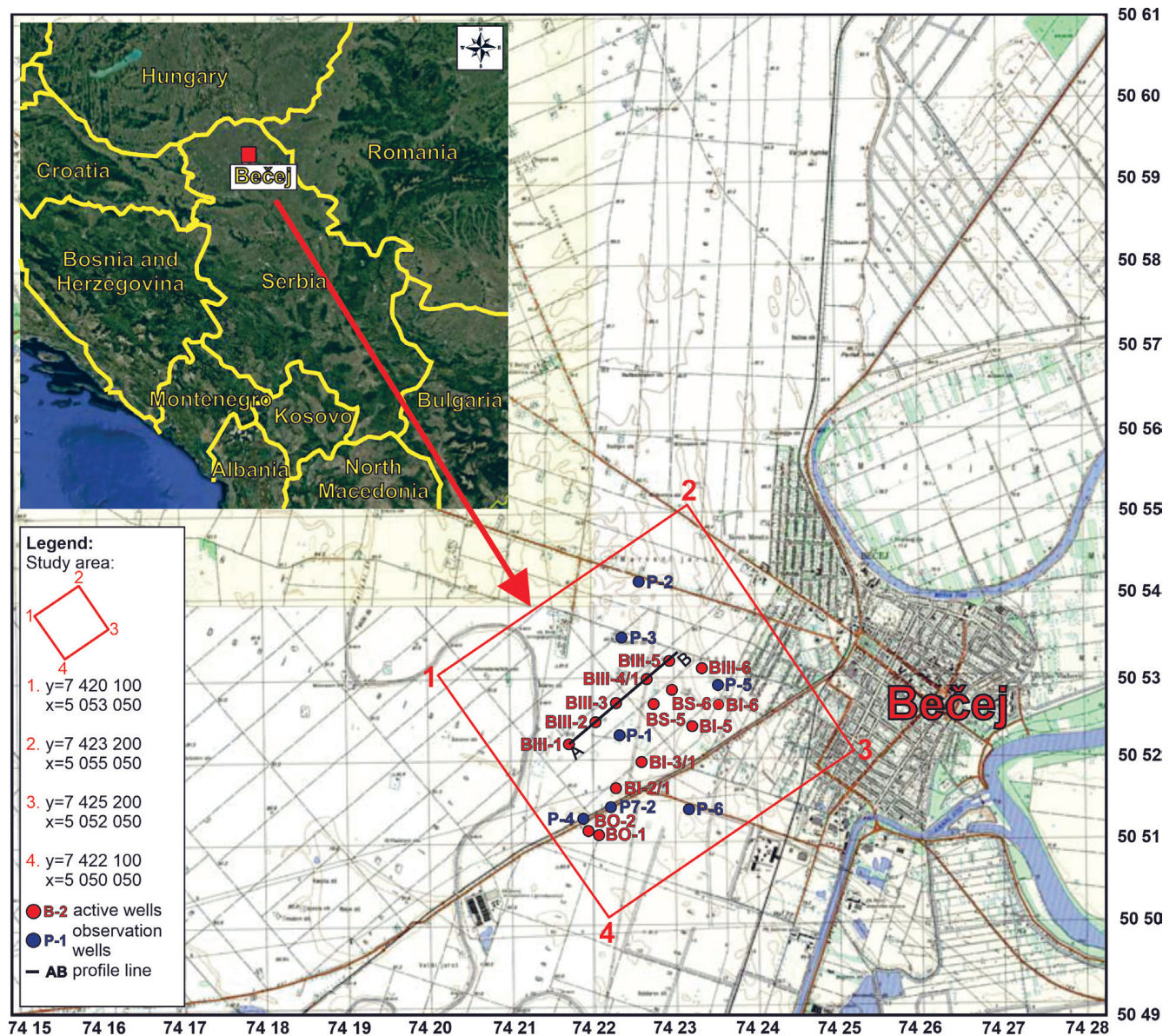

Figure 1. Location map of the study area.

Ratković et al.: Imputing missing data using grey system theory and the biplot method to forecast groundwater levels and yields   237

Geologia Croatica

Table 1. Well designations and tapped aquifers.

Active wells	Tapped aquifers
B0-1	second, third
B0-2	first, second
BI-2/1	second, third
BI-3/1	first, second, third
BI-5	second, third
BI-6	second, third
BIII-1	first, second, third
BIII-2	first, second, third
BIII-3	first, second, third
BIII-4/1	first, second, third
BIII-5	third
BIII-6	second, third
BS-5	first, second, third
BS-6	first, second, third

Table 1. includes the designations of the wells and shows the aquifers they capture.

The extended study area is located within the Pannonian Plain and is composed of Precambrian/Palaeozoic, Mesozoic, Mi-ocene, Pliocene and Quaternary rocks. The oldest rocks are crystalline schists and Precambrian/Palaeozoic granites. The youngest rocks are Quaternary in age, with alternating lithological units (gravels and coarse sands underlie sandy silts, silts and clays) (MALEŠEVIĆ, 1984; TERZIN, 1994). Vertically, there are two types of aquifers down to a depth of about 130 m: an unconfined intergranular aquifer in the alluvial and terrace deposits (from the land surface to roughly 30 m – aquifer I) and three confined intergranular aquifers in deeper geological formations (from 60 m to 130 m – aquifers II, III and IV). Bečej GS taps the deeper aquifers of Quaternary age: the first is at a depth of 60 to 80 m, the second at 83 to 100 m, and the third from 110 to 130 m (Fig. 2).

The unconfined aquifer is composed of sands in the upper part and fine-grained, partly clayey sands in the lower part. The bottom of this layer comprises sandy clay. The three deeper aquifers are composed of fine- to medium-grained sands, with some gravel. The sands include a pelitic component, which hinders production/causes well ageing. The aquifers are separated by relatively thin layers of clay, sandy clay and clayey sand. Their thickness ranges from 5 to 20 m. These three aquifers constitute a hydrogeologic complex, which evolved during the Lower and Middle Pleistocene. On a regional scale, they are considered as a single, confined aquifer (MALEŠEVIĆ, 1984; TERZIN, 1994).
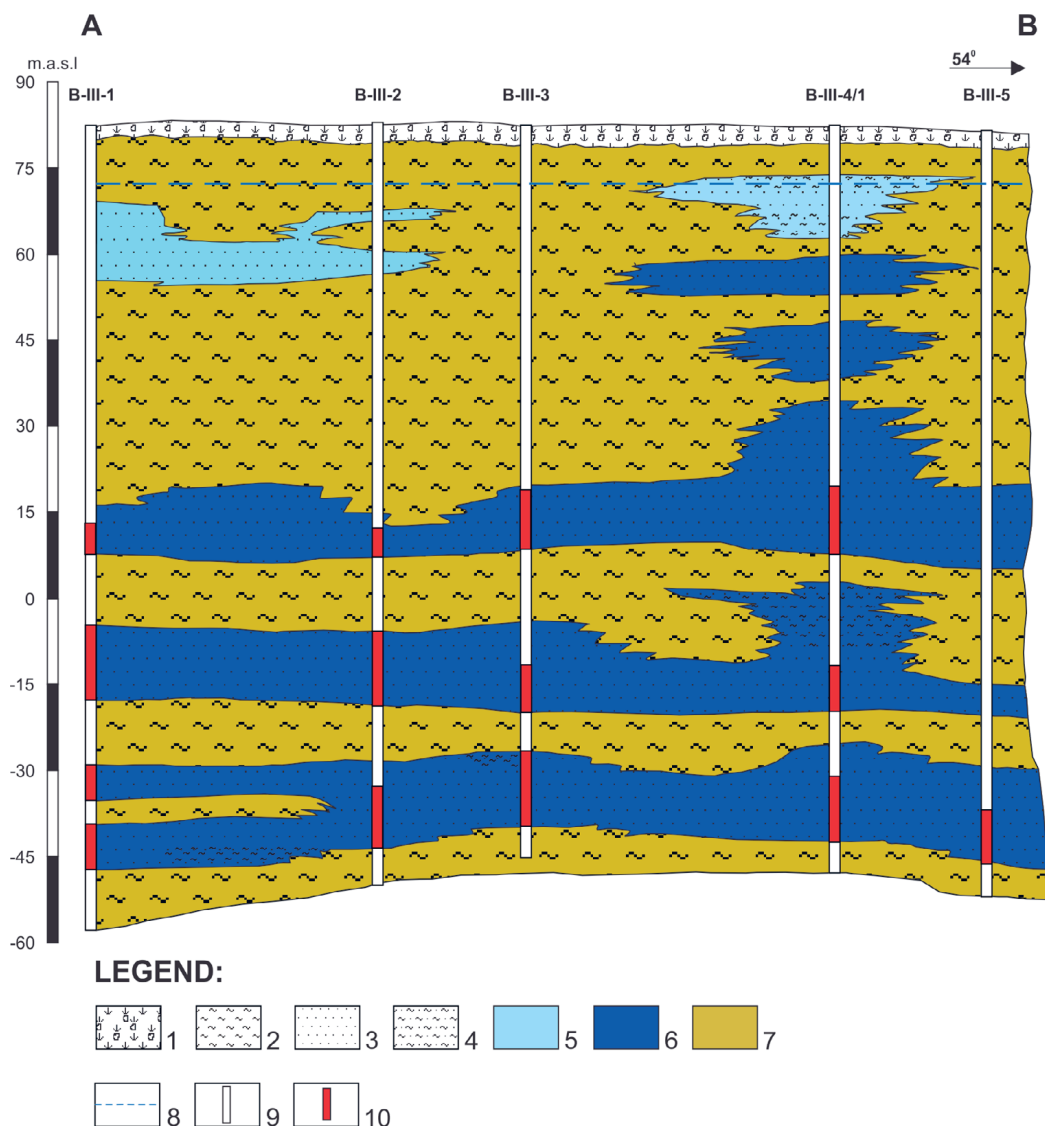

Figure 2. A hydrogeological section through the Bečej Groundwater Source. Legend: 1. surface decomposed material; 2. clay; 3. sand; 4. clayey sand; 5. unconfined aquifer; 6. confined aquifer; 7. semi-permeable sediments; 8. confined aquifer head; 9. well; and 10. position of well screen.

Geologia Croatica

3. METHODOLOGY – THE FORECASTING MODEL BUILDING PROCESS

3.1. Problem statement

The study period involving groundwater levels and groundwater abstraction rates at Bečej GS is from 1 October 1980 to 1 May 2010. However, piezometric head was not monitored from 1 February 1989 to 1 June 1998 and total source yield from 1 October 1980 to 1 June 1998. The monitoring network comprised seven observation wells (P1, P2, P3, P4, P5, P6, P7/2), at depths from 120 m to 150 and 14 pumping wells (B0-1, B0-2, BI-2/1, BI-3/1, BI-5, BI-6, BIII-1, BIII-2, BIII-3, BIII-4/1, BIII-5, BIII-6, BS-5, BS-6), at which total groundwater source capacity was gauged (Fig. 3). Based on the data, groundwater levels were observed much more frequently than the total source yield. These two parameters were recorded periodically, but slightly more often in 2006 to assess the water balance and determine groundwater reserves. The present research considers mean monthly values of the parameters, which were the only available data (Fig. 3).

As part of the reported research, a methodology was developed to reconstruct the missing data (10-year period) due to the lack of continuous monitoring. It is shown in Fig. 4 in the form of an algorithm.

In the first step, input data comprising values observed during the course of monitoring were analyzed. The objective of the proposed methodology was to reconstruct the missing data applying the methods indicated in the algorithm. The *AR model* was used to reconstruct the dates (observation positions) on which piezometric head and source yield were not observed. The accuracy of the AR model was determined on the basis of AIC (Akaike information criterion), according to which the lowest value was assumed to be accurate. After the observation dates were reconstructed, the *grey method* was applied to arrive at piezometric head values for the period during which there was no monitoring. The input data for the GM were the original and reconstructed dates of observation and the observed piezometric heads. First the GM was tested against the (original) data observed during the study period. If a comparison of the original and reconstructed piezometric head data and MAPE criterion (Table 2) indicated that the accuracy of the model was satisfactory, the approach was acceptable for reconstructing the missing data. If not, the proposed methodology was not suitable for further reconstruction. Given that calculations and the MAPE criterion (Table 2) indicated that the accuracy of GM was satisfactory, it was applied to fill the data gaps. The above methods used to determine the missing observation dates and piezometric heads produced a completed matrix for the *biplot method*, which was applied to reconstruct the total source yields during the period of no observations. The accuracy of the total source yields derived by the biplot method was determined based on the condition $\dfrac{d}{\bar{y}} < 0.01$ (where:

$$d = \left[\left(\frac{1}{na} \right) \sum_{i=1}^{na} \left(x_i - x_i^A \right)^2 \right]^{\frac{1}{2}}, \quad \bar{y} = \left[\left(\frac{1}{N} \right) \sum_{i=1}^{n} \sum_{j=1}^{p} y_{ij}^2 \right]^{\frac{1}{2}} \text{). If the}$$
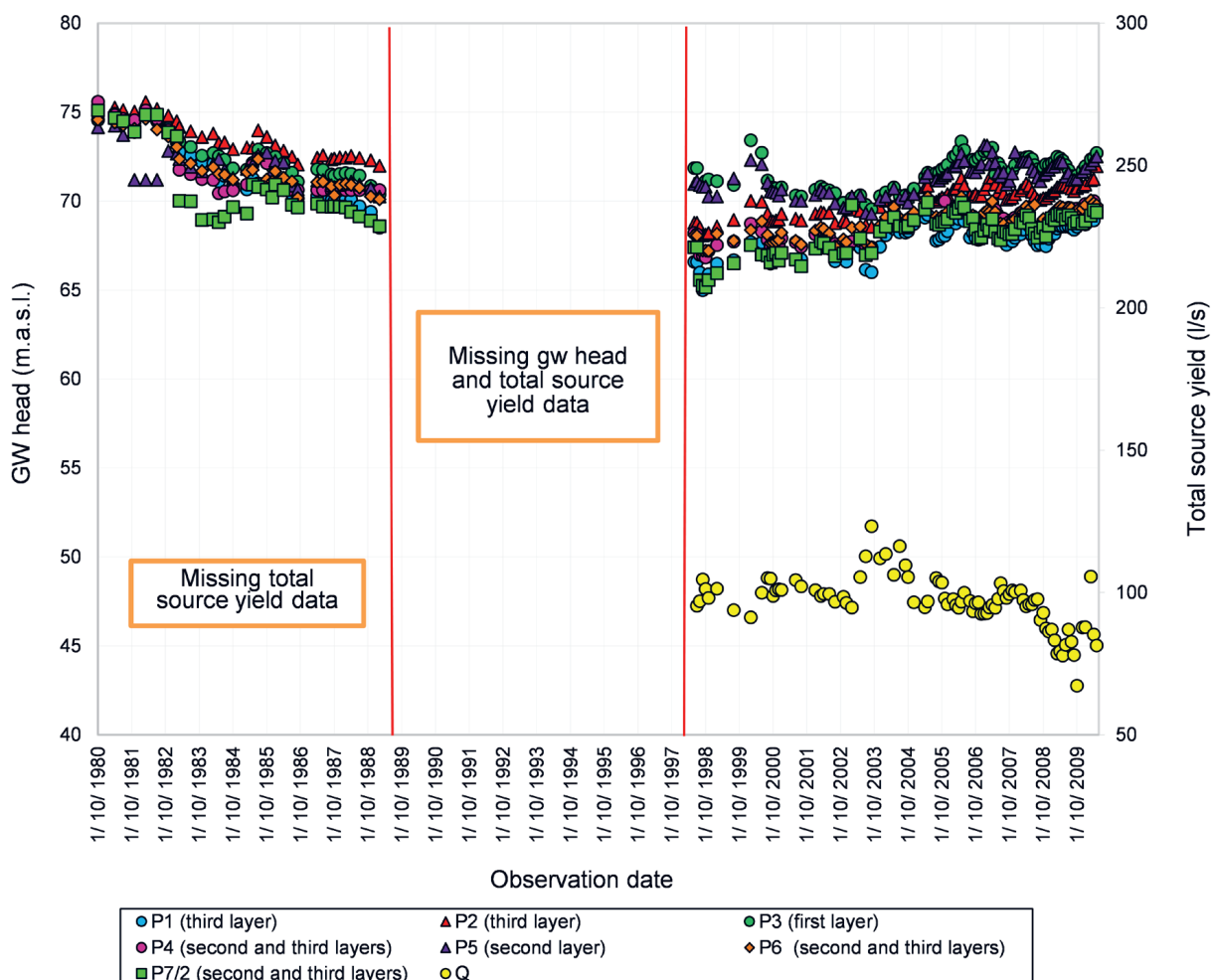

Figure 3. Piezometric head recorded at Bečej GS observation wells from 1 October 1980 to 1 May 2010 and total source yield.
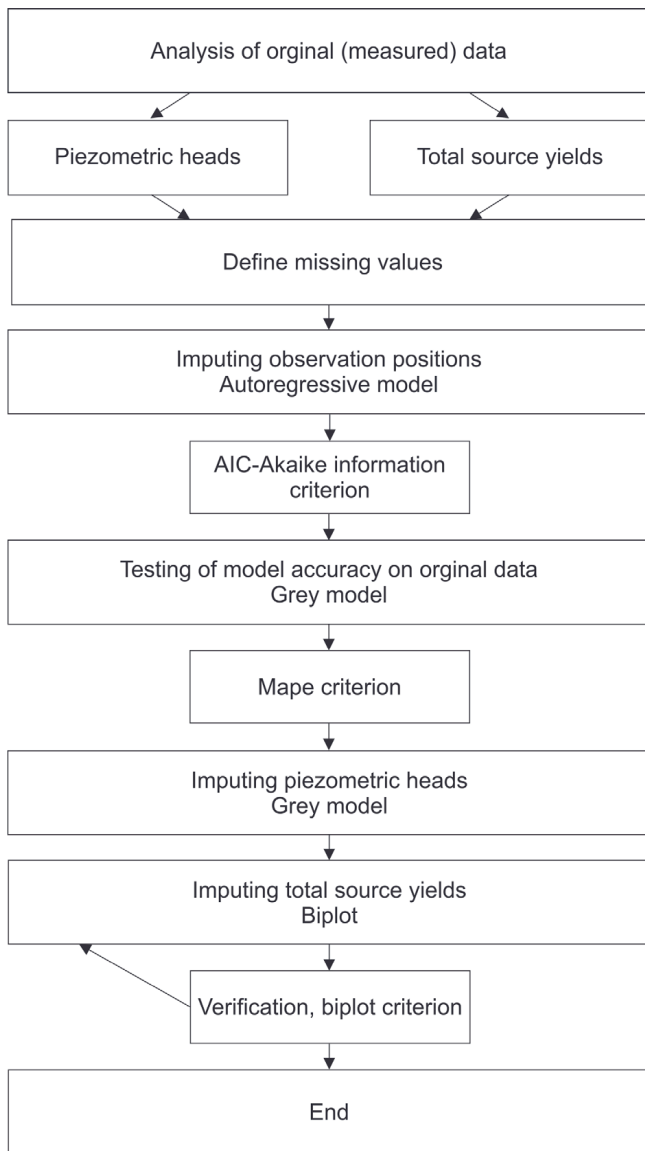
Figure 4. Operational research algorithm – the forecasting model building process.

solution was acceptable, the model was deemed suitable for reconstructing the total source yields. If not, it would have been necessary to look for other, more accurate models. All the calculations were made in Microsoft Excel (AR model, grey theory and biplot method) and MathLab (biplot method). The figures were produced in CorelDraw (Figs. 1, 2, 4) and Microsoft Excel (Figs. 3, 5, 6, 7, 8, 9).

A brief theoretical background of the models used in this research is provided below.

3.2. Autoregressive (AR) model

The AR model is commonly applied to process statistical time-series. The AR model was used for predicting missing observation positions. More specifically, it is used to forecast the values of current and future variables, based on the values of the same variable at different times. The variable is described as a function of its values in previous periods Z_{t-1}, Z_{t-2}.

The mathematical expression is (DAI et al., 2015):

$$X_t = \sum_{i=1}^{t-p} \alpha_i X_i + Z_t \qquad (1)$$

where p is the AR model order and Z_t is the constant.

In order to construct the AR model, first its order (p) needs to be determined. The principle is as follows: for the original (observed) data $(x_1, x_2,...,x_n)$ of order p, the AR model is:

$$X_t = \sum_{i=1}^{p} \alpha X_{t-1} + Z_t \qquad (2)$$

If the original time-series is stationary, the AR process is applied directly to the original dataset. If not, the time-series needs to be transformed into a stationary time-series (sequence). The common transformation approach is first-order differentiation, as was required in the present case.

First-order differentiation is implemented as:

$$\Delta X_t = X_t - X_{t-1} \qquad (3)$$

Then the AR model can be applied to the differentiated time-series. The ultimate forecasted values are derived from:

$$\hat{X}_t = X_{t-1} + \Delta \hat{X}_t \qquad (4)$$

where $\Delta \hat{X}_t$ is the value resulting from the AR process.

The AR model order is defined (determined) applying AIC. AIC is calculated as (CHAKRABORTY et al., 2009):

$$AIC = \ln\left(\frac{\sum_{i=1}^{N} \hat{u}_i^2}{N}\right) + \frac{2p}{N} \qquad (5)$$

where N is the length of the sample, p is the number of estimated coefficients (p-order, AR model order) and $\hat{u}_i$ are the residuals. The model where the AIC is the lowest is selected as optimal (most suitable).

Using the AR model, the observation positions (dates) are reconstructed and then used as input data for the GM.

3.3. Grey model (GM)

The theory provides an effective solution for the data uncertainty and discretization problem. GMs predict the future values of a time-series, based solely on a set of observed data (recorded, original values in the present case). According to the grey system theory, a general grey model GM(h,N) is the GM in which h is the order of the differential equation and N is the number of parameters.

GM(h,N) is defined by the following differential equation (YU et al., 2001):

$$\sum_{i=0}^{h} a_i \frac{d^{(i)} x_l^{(1)}(k)}{dt^{(i)}} = \sum_{j=2}^{N} b_j x_j^{(1)}(k), k = 1, 2, ..., n \qquad (6)$$

where a_i and b_j are the determined coefficients, $x_l^{(1)}(k)$ is the sequence of the main factor, $x_j^{(1)}(k)$ are the sequences of the influential factors, and k is the time sequence variable.

A single-variable first-order grey model GM(1,1) was used in the present research because there was a single time-series of observations.

The original time-series was represented as (ZHAO-HUI & JING, 2016):

$$X^{(0)} = \left[x^{(0)}(t_1), x^{(0)}(t_2),..., x^{(0)}(t_j) \right], j = 1,..., n, \text{ if:}$$

$$\Delta t_j = t_j - t_{j-1} \neq \text{const}$$

Before proceeding with the GM equation, the accumulating generation operator (AGO) will be defined. AGOs are used in GM to reduce data randomness.

The first-order AGO for $X^{(0)}$ is defined as (ZHAO-HUI & JING, 2016):

$$X^{(1)}(t_j) = X^{(1)}(t_{j-1}) + X^{(0)}(t_j) * \Delta t_j \qquad (7)$$

where $x^{(1)}(t_j) = \sum_{j=1}^{n} x^{(0)}(t_j)$.

The sequence of the generated averages of the neighboring values of sequence $X^{(1)}(t_j)$ is
$Z^{(1)}(t_j) = \left[z^{(1)}(1), z^{(1)}(2), ..., z^{(1)}(t_j) \right]$, where:

$$z^{(1)}(t_j) = (\tfrac{1}{2})(x^{(1)}(t_j) + x^{(1)}(t_j - 1)), \ t_j = 1, 2, ..., n \quad (8)$$

The grey differentiated first-order equation GM(1,1) is (ZHAO-HUI & JING, 2016):

$$\frac{dx^{(1)}(t)}{dt} = ax^{(1)}(t) + b \qquad (9)$$

The first term $x^{(1)}(t_1)$ of the sequence $x^{(1)}(t_j)(j = 1, 2, ..., n)$ is used as the initial condition of the grey differentiated equation.

The solution to Eq. (9) is a time-dependent function:

$$x^{(1)}(t) = e^{at} x^{(1)}(t_1) + a^{-1}(e^{at} - I)b \qquad (10)$$

To distinguish between a and b, Eq. (9) is discretized. In other words, the difference in grey derivative I of interval $\left[t_{j-1}, t_j \right]$ yields $x^{(0)}(t_j) = az^{(1)}(t_j) + b$

and thus:

$x^{(0)}(t_j) =$
$a(z^{(1)}(t_j) + z^{(1)}(t_{j-1})) + b \quad (i = 1, 2, ..., n; k = 2, 3, ..., m) \quad (11)$

Coefficients a and b are obtained by the least squares method:

$$(a, b)^T = (B^T B)^{-1} B^T Y \qquad (12)$$

where:

$$B^{(1)}(t_j) = \begin{bmatrix} \frac{1}{2}(x_1^{(1)}(t_2) + x_1^{(1)}(t_1)) & 1 \\ \frac{1}{2}(x_1^{(1)}(t_3) + x_1^{(1)}(t_2)) & 1 \\ ... & 1 \\ \frac{1}{2}(x_1^{(1)}(t_j) + x_1^{(1)}(t_{j-1})) & 1 \end{bmatrix} \qquad (13)$$

$$Y = \left[x^{(0)}(t_2), x^{(0)}(t_3), ..., x^{(0)}(t_j) \right]^T$$

The time sequence response formula, that is, the solution of (9) is:

$$x^{(1)}(t_j) = e^{a(t_j - t_1)} x^{(1)}(t_1) + a^{-1}(e^{a(t_j - t_1)} I) * b \ \ j = 1, 2, ..., n \ \ (14)$$

The fitting forecast value of xj(0) can be obtained by the inverse AGO, as follows (LIU & LUO, 2010):

$$\hat{x}^{(0)}(t_j) = \frac{x^{(1)}(t_j) - x^{(1)}(t_{j-1})}{\Delta t_j}, j = 2, ..., n \qquad (15)$$

Table 2. Accuracy according to MAPE (%).

MAPE (%)	Forecast (prediction) accuracy
< 1	High
1-5	Good
5-10	Acceptable
>10	Unacceptable (low)

The accuracy of the model is represented by the mean absolute percentage error (MAPE) (ZHAO-HUI & JING, 2016):

$$MAPE = \frac{1}{n} \sum_{j=1}^{n} \left| e(t_j) \right| \qquad (16)$$

where:

$$e(t_j) = \frac{\hat{x}^{(0)}(t_j) - x^{(0)}(t_j)}{x^{(0)}(t_j)} \qquad (17)$$

In the present case, MAPE was used to assess model accuracy. The model accuracy criterion is shown in Table 2 (WANG & PHAN, 2016):

The GM (1,1) was used to reconstruct piezometric head data during the 10-year period of discontinued observations (from 1 February 1989 to 1 June 1998, Fig. 3). The completed observation positions (dates) and piezometric heads represented input data for the biplot method.

3.4. Biplot method

The biplot method involves the imputing of arbitrary values to missing data in order to complete the matrix, and then calculating SVD using only two components (ARCINIEGAS-ALARCON et al., 2014). SVD is the basis of the biplot method. In the reported case study, the biplot method was used to forecast missing data on total source yields from 1 October 1980 to 1 June 1998.

The initial data for constructing the matrix included:
1. Original and reconstructed observation positions (dates) on which groundwater level and total source yield were observed);
2. Piezometric head data – original and reconstructed using the GM (1,1); and
3. Observed source yield data.

The constructed nxp matrix denoted by X comprised x_{ij} ($i=1,...,n$; $j=1,...,p$) elements, where the missing data were symbolized by x_{ij}^{aus}. In the first step, the missing values were imputed in their respective columns, to complete matrix X. In the specific case, observed average discharges (i.e. known values) were used to fill the data gaps.

The completed matrix X was standardized:

$$p_{ij} = \frac{(x_{ij} - m_j)}{s_j} \qquad (18)$$

where:

m_j is the mean deviation of the j-th column,
s_j is the standard deviation of the j-th column, and
p_{ij} are the standardized elements.

The matrix with p_{ij} elements was denoted by P.

The next step was to calculate SVD of matrix P. SVD of a standardized matrix is a fundamental and prominent approach in modern numerical linear algebra. The SVD algorithm represents an approximation of the main matrix by smaller matrices (MEI-

JAARD, 1993). According to this theorem, any matrix A (mxn) can be decomposed into three matrices, as:

$$A = ULV^t \qquad (19)$$

U is the orthogonal mxm matrix,

V is the orthogonal nxn matrix,

L is the diagonal mxn matrix with positive elements on the main diagonal, referred to as singular values of A, and all the other terms are equal to zero.

The singular values are generally sorted in descending order, on the main diagonal L.

Taking into account the first two components, the resulting equation (expression) is:

$$p_{ij} = \frac{(x_{ij} - m_j)}{s_j} = \sum_{k=1}^{2} \lambda_k \alpha_{ik} \gamma_{jk} + \varepsilon_{ij} \qquad (20)$$

λ_k are the singular values,

α_{ik} are the eigenvectors of the rows,

γ_{jk} are the eigenvectors of the columns, and

ε_{ij} is the error of row i in column j.

The new matrix, composed of $p_{ij}^{(2)}$ elements, was denoted by $P^{(2)}$. All the $p_{ij}^{(2)}$ elements in $P^{(2)}$ were re-assigned their original values,

$\hat{x}_{ij}^{(2)} = m_j + s_j p_{ij}^{(2)}$. This resulted in a new matrix – $X^{(2)}_{(n \times p)}$. The x_{ij}^{aus} missing in the original matrix X were imputed with the corresponding values of $\hat{x}_{ij}^{(2)}$ in $X^{(2)}$. Filling of data gaps was followed by iterations (back to matrix standardization, repeating the entire process) until the specified condition ($\frac{d}{\bar{y}} < 0.01$) was fulfilled and the imputed missing values stabilized. The iterations then continued until the following condition was met (ARCINIEGAS-ALARCON et al., 2014):

$$\frac{d}{\bar{y}} < 0.01 \qquad (21)$$

where:

$$d = \left[\left(\frac{1}{na} \right) \sum_{i=1}^{na} (x_i - x_i^A)^2 \right]^{\frac{1}{2}} \qquad (22)$$

and

$$\bar{y} = \left[\left(\frac{1}{N} \right) \sum_{i=1}^{n} \sum_{j=1}^{p} y_{ij}^2 \right]^{\frac{1}{2}} \qquad (23)$$

– na is the total number of values missing in matrix X,
– x_i is the value forecasted for the i-th row of missing values in the active iteration,
– x_i^A is the value forecasted for the i-th missing values in the previous iteration,
– y_{ij} are the original values (not missing) in the i-th row and j-th column, and
– N is the total number of considered values.

4. RESULTS AND DISCUSSION – APPLICATION OF THE METHODOLOGY TO A REAL CASE STUDY

Before the missing data were reconstructed, the hydrogeological system was analyzed based on data collected in the last 30 years of operation to determine the groundwater dynamics.

The unconfined intergranular aquifer (aquifer I) is relatively thin and the values of its hydraulic parameters are low, so the groundwater reserves are modest. This aquifer is recharged by precipitation and via a good hydraulic link with a river. The aquifer is drained by the pumping wells and by migration into the deeper aquifers, given the semi-permeable floor of aquifer I.

Recharge of the confined intergranular aquifers (II, III and IV) is highly complex because of the large distances involved, so the recharge process can only be assumed based on the geological and hydrogeological features of the terrain and long-term piezometric levels of the groundwater. In view of the fact that the spread of the aquifers is regional, recharge occurs over a large area where their overlying strata are dominated by sands or in zones where the aquifers are directly linked to the land surface or rivers. Infiltration of precipitation into the confined aquifers has not been confirmed or occurs on a small scale. It is assumed that recharge also takes place through groundwater migration from the first, unconfined aquifer, given that the strata overlying aquifer II are composed of sands, clayey sands and sandy clays, which allow partial groundwater infiltration from aquifer I. This infiltration rate is very slow. The confined aquifers are drained solely by the studied pumping wells and neighbouring groundwater sources.

All the observation wells registered identical piezometric head fluctuation trends – lower in the summer and early autumn months due to higher rates of groundwater abstraction. Given the amount of water withdrawn from these aquifers, this decrease indicates that recharge conditions are satisfactory. The total average monthly yield of the groundwater source suggests that the amount of abstracted groundwater has not changed significantly compared to 10 years ago. The recorded piezometric level ranged from a minimum of 65 m above sea level to a maximum of 76 m.a.s.l., while the total source yield was from 68 l/s to 124 l/s.

The missing data reconstruction results are described below. Records of groundwater abstraction rates since the groundwater source was commissioned (a period of 30 years, Fig. 3 provided input data for applying the GM (1,1) and biplot method to Bečej GS. The missing data were reconstructed according to the algorithm shown in Fig. 4, first the observation dates using the AR model, then the groundwater levels resulting from GM (1,1), and finally the total source yield applying the biplot method. The proposed methodology with regard to the period of no observations is described below.

4.1. Imputing of missing observation positions based on the AR model

The AR model was used to forecast observation data (dates) of piezometric head and total source yield. Observation date is the date (day, month and year) of measurement. For AR modeling
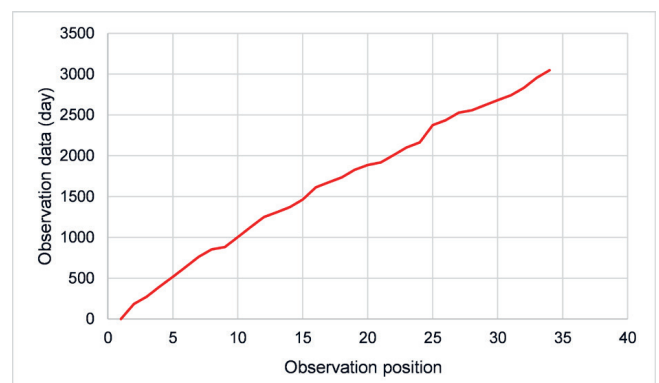

Figure 5. Original piezometric head observation positions at all observation wells.

purposes, these dates were converted into cumulative days (Fig. 3 shows the dates of original measurement). Each observation date is identified as an observation position in order to define the total number of measured and later reconstructed values (125 measured and 38 reconstructed).

Given that the original time-series was transient (Fig. 5) and in view of the growing trend resulting from Eq. (3), first-order differentiation (Fig. 6) was needed in order to apply the AR model. Time-series that exhibit a trend are referred to as non-stationary.

Equation (1) was applied to the stationary time-series, the values for the AR models of the second, third and fourth orders were calculated, and then the AIC criterion was used to assess the most accurate solution. AIC was applied to second-, third- and fourth-order AR (AR2, AR3 and AR4). The results based on Eq. (5) are shown in Table 3.

The lowest AIC value was adopted (i.e. second-order AR, AR2).

Figure 7. shows the original dates (in red) from 1 October 1980 to 1 February 1989, based on which the missing dates (shown in blue) were reconstructed, along with subsequent original dates (1 June 1998 to 1 May 2010).

Table 4. is a numerical representation of the forecasted observation positions based on second-order AR. The dates shown in red were reconstructed (there were no actual piezometric head and source yield observations).

The MAPE of AR2 was 2.26 %. In view of the large number of missing observation dates, this error was deemed acceptable. Also, AR2 was adopted according to AIC.

4.2. Imputing of missing piezometric head data based on the GM

The input data used in this research were considered to be a grey system, with some known and some unknown information, so

Figure 6. Stationary time-series of original observation positions at all observation wells.
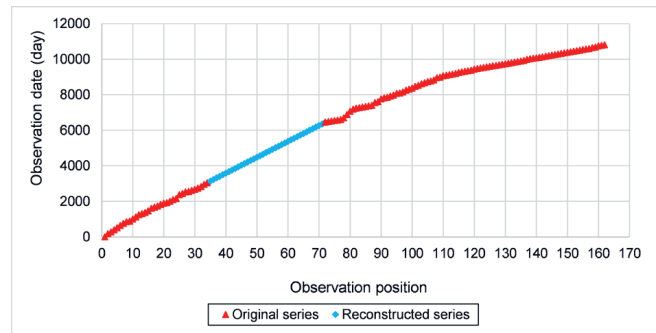
Figure 7. Original and reconstructed observation positions according to second-order AR.

that it could be analyzed applying the grey system theory. The first step was to analyze the input data from 1 October 1980 to 1 May 2010, and the second step to apply a GM (1,1). The GM (1,1) was used to reconstruct piezometric head values for the period during which the observation wells (P1, P2, P3, P4, P5, P6, P7-2) were not monitored. The first step was to verify the accuracy of the GM (1,1) using the observed data.

Table 3. AIC values for AR2, AR3 and AR4.

AIC(AR2)	AIC(AR3)	AIC(AR4)
7.43	7.44	7.44

Table 4. Reconstructed observation positions based on second-order AR.

Observation position	Day	Observation position	Day	Observation position	Day	Observation position	Day
1	1	19	1827	37	3315	55	4924
2	183	20	1888	38	3405	56	5013
3	274	21	1919	39	3494	57	5102
4	397	22	2009	40	3583	58	5192
5	517	23	2100	41	3673	59	5281
6	639	24	2162	42	3762	60	5371
7	762	25	2374	43	3851	61	5460
8	854	26	2435	44	3941	62	5549
9	882	27	2527	45	4030	63	5639
10	1004	28	2557	46	4119	64	5728
11	1127	29	2618	47	4209	65	5817
12	1248	30	2680	48	4298	66	5907
13	1309	31	2740	49	4388	67	5996
14	1370	32	2831	50	4477	68	6085
15	1462	33	2954	51	4566	69	6175
16	1613	34	3046	52	4656	70	6264
17	1674	35	3136	53	4745	71	6353
18	1735	36	3226	54	4834	72	6443

Table 5. Observed (original) and reconstructed values of piezometric head at observation well P1 and absolute error.

Day	P1 Original time-series	P1 Recon. time-series	Abs. error	Day	P1 Original time-series	P1 Recon. time-series	Abs. error	Day	P1 Original time-series	P1 Recon. time-series	Abs. error
1	74.57	74.57	0.00	7063	67.69	68.76	1.58	9497	67.94	67.72	0.33
183	74.74	71.69	4.08	7184	67.68	68.70	1.50	9528	67.82	67.71	0.17
274	74.59	71.63	3.97	7245	67.02	68.66	2.44	9558	67.89	67.69	0.29
397	74.41	71.58	3.80	7276	66.49	68.64	3.23	9589	68.41	67.68	1.07
517	75.01	71.53	4.64	7306	66.59	68.63	3.06	9620	68.64	67.67	1.42
639	74.50	71.48	4.05	7337	66.63	68.61	2.98	9648	68.52	67.66	1.26
762	74.11	71.43	3.62	7367	66.68	68.60	2.88	9679	68.98	67.64	1.94
854	73.28	71.38	2.59	7398	67.07	68.59	2.26	9709	68.60	67.63	1.41
882	72.77	71.35	1.95	7549	66.71	68.55	2.76	9740	68.20	67.62	0.85
1004	72.55	71.32	1.69	7610	66.74	68.51	2.65	9770	68.29	67.61	1.00
1127	72.18	71.27	1.26	7763	67.37	68.46	1.62	9801	68.09	67.59	0.73
1248	72.13	71.22	1.27	7822	67.52	68.42	1.33	9832	67.54	67.58	0.06
1309	71.11	71.18	0.09	7853	67.50	68.40	1.33	9862	67.83	67.57	0.38
1370	71.42	71.15	0.38	7914	67.34	68.38	1.54	9893	67.74	67.56	0.27
1462	70.78	71.12	0.48	7975	66.62	68.35	2.60	9923	67.94	67.54	0.58
1613	70.65	71.06	0.59	8067	66.91	68.32	2.11	9985	68.14	67.53	0.90
1674	71.27	71.02	0.35	8097	66.61	68.30	2.53	10014	68.39	67.51	1.29
1735	70.99	70.99	0.00	8159	67.58	68.28	1.03	10045	68.79	67.49	1.88
1827	70.79	70.96	0.24	8248	67.38	68.25	1.29	10075	68.60	67.48	1.63
1888	70.44	70.93	0.69	8309	66.15	68.22	3.12	10106	68.06	67.47	0.87
1919	70.91	70.91	0.01	8371	66.00	68.19	3.32	10136	67.74	67.46	0.42
2009	70.68	70.88	0.28	8462	67.43	68.16	1.08	10167	67.52	67.44	0.11
2100	70.39	70.84	0.64	8524	68.08	68.13	0.07	10198	67.59	67.43	0.23
2162	70.15	70.81	0.94	8614	68.44	68.10	0.50	10228	67.87	67.42	0.66
2374	70.30	70.75	0.64	8675	68.24	68.06	0.26	10259	67.47	67.41	0.09
2435	70.35	70.69	0.48	8737	68.23	68.04	0.28	10289	67.89	67.39	0.73
2527	70.03	70.66	0.90	8767	68.32	68.02	0.44	10320	68.08	67.38	1.03
2557	70.00	70.63	0.90	8828	68.72	68.00	1.05	10351	68.17	67.37	1.18
2618	70.09	70.61	0.75	8949	69.08	67.96	1.62	10379	68.46	67.36	1.61
2680	70.11	70.59	0.68	8979	69.23	67.93	1.87	10410	68.59	67.34	1.82
2740	70.09	70.56	0.67	9071	67.77	67.91	0.20	10440	68.57	67.33	1.80
2831	69.71	70.53	1.17	9102	67.84	67.88	0.06	10471	68.57	67.32	1.82
2954	69.41	70.48	1.54	9132	68.01	67.87	0.21	10501	68.57	67.31	1.84
3046	68.51	70.44	2.81	9163	68.06	67.86	0.30	10532	68.49	67.30	1.74
6453	66.59	69.69	4.66	9193	68.32	67.84	0.70	10563	68.39	67.28	1.62
6483	66.59	68.97	3.57	9255	68.85	67.82	1.49	10593	68.56	67.27	1.88
6514	65.99	68.96	4.49	9283	68.75	67.81	1.37	10654	68.74	67.25	2.17
6545	64.99	68.94	6.08	9314	68.99	67.79	1.73	10685	68.69	67.23	2.12
6575	65.54	68.93	5.17	9344	70.41	67.78	3.73	10744	68.99	67.21	2.57
6606	65.89	68.92	4.60	9375	68.89	67.77	1.63	10775	68.92	67.20	2.50
6698	66.49	68.89	3.61	9436	67.95	67.75	0.29	10805	69.49	67.18	3.32
6879	66.69	68.83	3.22	9467	68.22	67.73	0.72				

Equations (7) through (15), described in the theoretical section, were applied in that order to the original time-series of observed piezometric head data. This yielded the predicted piezometric head values. The objective was to check the accuracy and applicability of the GM to the case study or, in other words, to the original measured data (Table 5). Then Eq. (16) was applied to the reconstructed time-series to check the calculation error and Table 2. was used to assess whether the GM was applicable to the original time-series and if it yielded acceptable solutions.

Table 5. shows the observed (original) and reconstructed piezometric heads at observation well P1. It is apparent that the difference between the observed and reconstructed values was acceptable according to the criterion from Table 2 and that the model could be used for the discontinued period from 2 January 1989 to 6 January 1998.

The errors are shown in Table 6.

The observed and reconstructed piezometric head values agreed rather well, given the ten-year period of discontinued

Geologia Croatica

Table 6. Reconstructed piezometric head errors at all observation wells.

	Max. abs. error	MAPE (%)
P1	6.08	1.60
P2	4.42	1.34
P3	3.31	1.16
P4	4.51	1.40
P5	3.61	1.02
P6	3.81	1.27
P7-2	5.83	1.66

groundwater monitoring (Table 5). Based on the criterion shown in Table 2, the conclusion was that the error was acceptable in view of the fact that the accuracy of the reconstructed values was *high,* so the solution was deemed acceptable.

4.3. Imputing of missing source yield data based on the biplot method

The biplot method was used to reconstruct the missing source yield data, from 1 October 1980 to 1 July 1998. The input data in this case included original and reconstructed:
  – observation dates,
  – piezometric heads and
  – observed source yield data.

The first step, before the biplot method, was to complete matrix X for calculating the missing total source yield values. The averages of the observed source yields were used for the missing values, x_{ij}^{aus}. The completed matrix included the observation dates (original and reconstructed), piezometric heads (original and reconstructed), observed source yields and average source yields (in place of those that needed to be reconstructed). Matrix X was than standardized applying Eq. 18, resulting in standardized matrix P. This was followed by the SVD of matrix P (in Mat-Lab), based on Eq. 19. Then Eq. 20 was applied, resulting in matrix $P^{(2)}$. After applying $\hat{x}_{ij}^{(2)} = m_j + s_j p_{ij}^{(2)}$, all the elements of that matrix reverted to original values in matrix $X^{(2)}_{(nxp)}$. The missing x_{ij}^{aus} in the original matrix X were completed with the corresponding values of $\hat{x}_{ij}^{(2)}$ iz $X^{(2)}$. This was followed by iterations, back to matrix standardization, and the entire process was repeated. Iterations lasted until the condition $\frac{d}{\bar{y}} < 0.01$ was fulfilled (ARCINIEGAS-ALARCON et al., 2014), which occurred in the second iteration (Table 7).

Table 8 and Figure 8 show the observed (original) and reconstructed source yields.

Table 8. Observed (original) and reconstructed source yields – iteration 2.

Day	Q Original time-series	Q Reconst. time-series	Day	Q Original time-series	Q Reconst. time-series	Day	Q Original time-series	Q Reconst. series
6483	95.31	93.18	8614	106.28	106.08	9832	98.10	98.15
6514	96.99	95.85	8675	116.34	113.58	9862	99.40	98.39
6545	104.60	99.85	8737	109.64	109.02	9893	100.72	97.37
6575	101.23	97.75	8767	105.43	108.55	9923	100.29	93.04
6606	98.06	97.95	8828	96.55	101.54	9985	100.73	94.78
6698	101.41	100.84	8949	94.83	97.10	10014	97.35	93.02
6879	93.87	94.02	8979	96.86	97.34	10045	95.14	92.17
7063	91.33	86.23	9071	105.12	101.47	10075	95.83	92.19
7184	99.94	92.97	9102	103.90	100.80	10106	95.86	92.81
7245	105.14	101.06	9132	103.47	99.98	10136	97.43	94.89
7276	104.93	102.58	9163	97.93	96.67	10167	97.71	95.77
7306	98.80	99.41	9193	95.82	94.82	10198	90.37	92.84
7337	100.77	100.84	9255	97.66	94.33	10228	92.88	92.99
7367	101.13	101.91	9283	95.60	92.87	10259	87.41	89.06
7398	100.97	100.85	9314	94.68	91.59	10289	86.34	88.36
7549	104.38	105.04	9344	96.76	91.21	10320	86.97	88.42
7610	102.22	103.67	9375	99.93	94.22	10351	83.21	86.20
7763	100.82	102.24	9436	97.11	94.61	10379	78.58	82.91
7822	98.87	100.19	9467	93.47	92.56	10410	79.41	83.37
7853	99.52	101.09	9497	96.28	93.79	10440	77.86	82.76
7914	99.53	101.06	9528	96.59	92.61	10471	81.66	87.75
7975	96.76	100.87	9558	92.51	89.59	10501	86.99	90.80
8067	98.54	103.96	9589	92.61	86.49	10532	82.75	90.24
8097	96.35	103.06	9620	92.85	86.91	10563	78.07	87.35
8159	94.78	100.60	9648	94.74	91.53	10593	67.28	81.33
8248	105.43	105.68	9679	95.66	90.18	10654	87.77	91.09
8309	112.72	112.49	9709	94.59	93.39	10685	87.84	90.26
8371	123.38	118.85	9740	97.82	95.18	10744	105.61	99.02
8462	112.00	110.48	9770	103.30	99.40	10775	85.27	86.95
8524	113.51	111.08	9801	100.54	97.68	10805	81.44	83.75

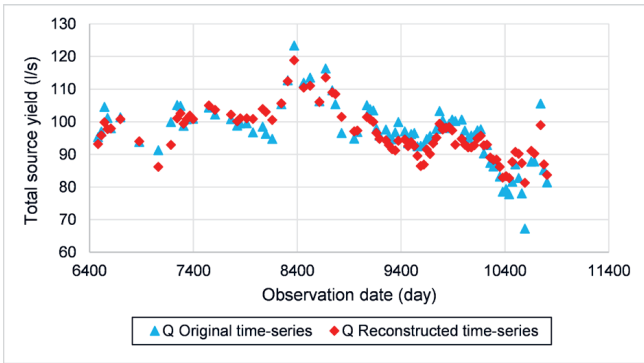**Figure 8.** Observed (original) and reconstructed source yield data (2nd iteration).

Table 7. Condition for ending iterations.

d	$\bar{y}$	$\dfrac{d}{\bar{y}}$
1.417	246.031	0.006

It is apparent in Figure 8. that the original and reconstructed source yield data agree rather well, over the entire study period (from 1 October 1980 to 1 May 2010).

Table 9. shows the errors in the 2nd iteration.

Figure 9. shows the Bečej GS observed and reconstructed piezometric head and source yield data.
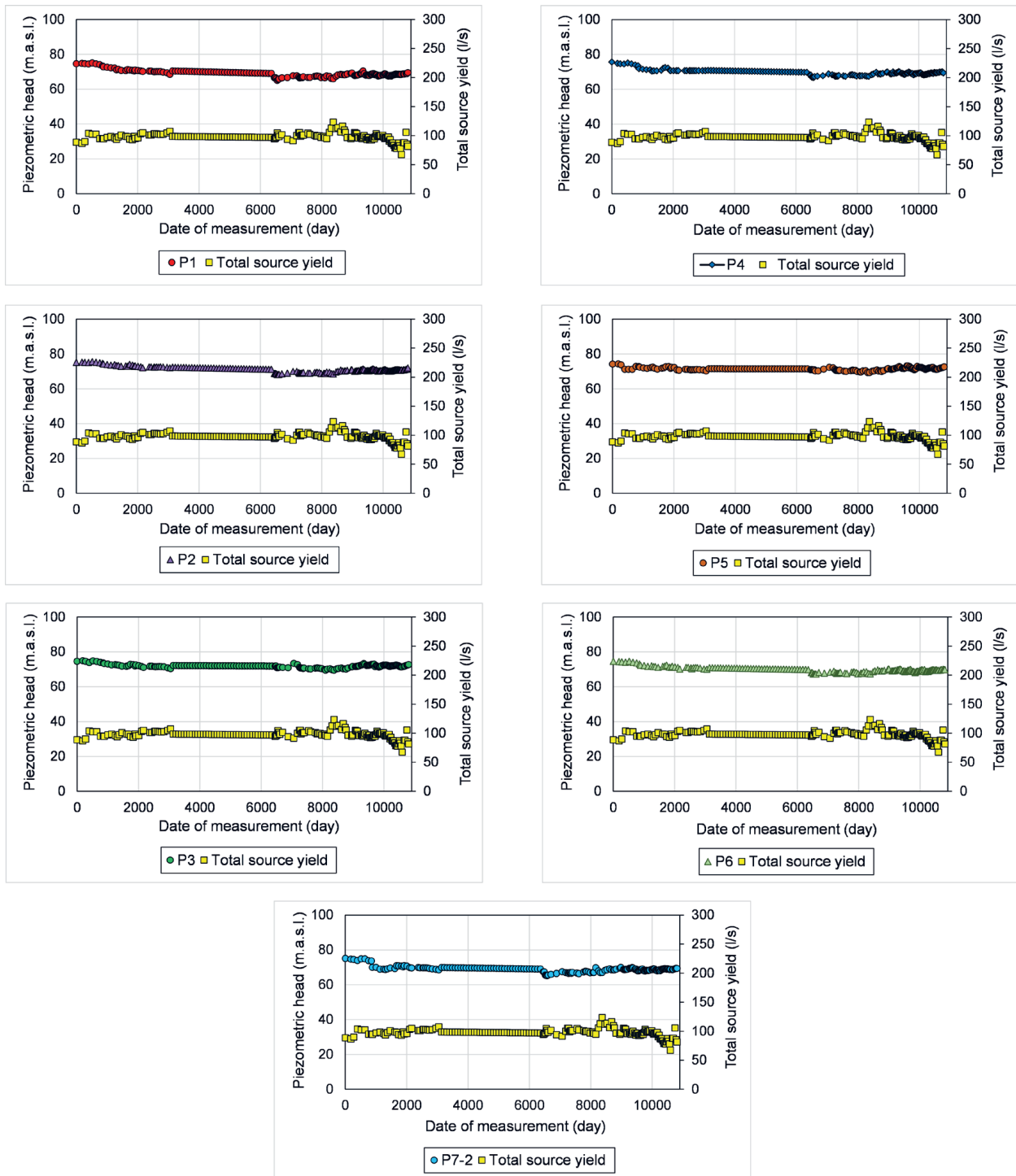
Figure 9. Original (observed) and reconstructed piezometric head and total source yield at all observation wells of Bečej GS.

Geologia Croatica

Table 9. Second iteration errors

Iteration 2 (%)	
Min. error	-7.78
Max. error	17.27
MAPE	3.32

Figure 9. shows that the trends of the piezometric head and source yield data are consistent with those registered *in situ*. The reason for the lack of fluctuations of the reconstructed values is attributable to the long period of time during which the considered parameters were not observed, as the applied methodology cannot produce more accurate solutions. However, the proposed methodology was highly effective in the study area, given that all the solutions were within the set criteria and the errors were in the predefined range. The models used in this research yielded satisfactory results and it is believed that they can be applied to fill data gaps in the study area (Bečej, Serbia). The developed methodology can also serve as a tool for upgrading other models (e.g. hydrodynamic) in groundwater management, when available input data are insufficient for this type of analysis. The level of error associated with the proposed methodology indicates that it can be applied to other groundwater sources and groundwater management systems.

5. CONCLUSION

An adequate groundwater source monitoring database is essential for efficient decision making in groundwater resources management. When such data are lacking, it is a challenge to provide sufficiently reliable groundwork for forecasting the future status of a groundwater source and the regime of the captured groundwater.

The paper described a methodology developed to reconstruct missing piezometric head and source yield data because monitoring of the Bečej groundwater source (Bečej GS) had been discontinued for a period of ten out of 30 years. The accuracy of the reconstructed values was first assessed in respect of the observed (original) data. MAPE showed that the applied models were highly accurate, indicating that the methodology could be applied to the entire time series of data observed during the study period.

Temporarily discontinued monitoring of piezometric head and source yield is one of the problems that had an adverse effect on the reliability of decisions relating to the feasibility of expanding the groundwater source in the future, in view of the potential of the area and the predicted efficiency of groundwater abstraction over a long period of time.

The proposed methodology effectively solves that problem. The implementation of the proposed algorithm, resulting in complete input data, improved source expansion solutions based on one of many methods for predicting groundwater dynamics. In the specific case, future research will focus on hydrodynamic modeling with reconstructed and filled data gaps for Bečej GS.

The proposed, verified methodology can be applied to groundwater abstraction systems, but also other systems, such as groundwater control (mines, hydraulic structures, agriculture, industrial zones, and urban areas), groundwater protection, and hydrometeorological, hydrological and similar systems.

The advantage of the proposed methodology is that it ensures sufficiently reliable reconstruction and/or forecasting of needed parameters, in cases that involve a small amount of observed data, or known quantities. In some instances it is relatively simple to implement.

AUTHOR CONTRIBUTIONS: All authors discussed the results and contributed to the final manuscript. J.R. with D.P., Z.G. and D.B. designed the methodology and wrote the final manuscript. Research plan, fieldwork and initial data analysis was conducted by J.R. D.B. provided critical feedback and helped shape the research methodology.

CONFLICTS OF INTEREST: The authors declare no conflicts of interest.

REFERENCES

ALVISI, S. & FRANCHINI, M. (2011): Fuzzy neural networks for water level and discharge forecasting with uncertainty.– Environmental Modelling & Software, 26, 523–537. doi: 10.1016/j.envsoft.2010.10.016

ARCINIEGAS-ALARCON, S., GARCIA-PENA, M., KRZANOWSKI, WJ. & DIAS, C.T.S. (2014): Imputing missing values in multi-environment trials using the singular value decomposition: An empirical comparison.– International journal of the faculty of agriculture and biology, 9, 54–70. doi: 10.1590/1984-70332016v16n2a13

CHAKRABORTY, S., DENIS, D.M. & SHERRING, A. (2009): Development of Time Series Autoregressive Model for prediction of rainfall and runoff in Kelo Watershed Chhattisgarh.– International Journal of Advances in Engineering Science and Technology, 2, 153–163.

CHITSAZAN, M., RAHMANI, G. & NEYAMADPOUR, A. (2015): Forecasting Groundwater Level by Artificial Neural Networks as an Alternative Approach to Groundwater Modeling.– Journal of the Geological Society of India, 85, 98–106. doi: 10.1007/s12594-015-0197-4

DAI, X., LIU, J. & ZHANG, H. (2015): Application of AR Model in the Analysis of Preearthquake Ionospheric Anomalies.– Mathematical Problems in Engineering, ID ID 157184. doi: 10.1155/2015/157184

DAOGONG, H.U., SHUREN, W.U. & CHENGXUAN, T.A.N. (1999): Grey Prediction of Seismic Activity in the Three Gorges Area of the Yangtze River.– Acta Geologica Sinica, 73, 294–299.

EMAMGHOLIZADEH, S., MOSLEMI, K. & KARAMI, G. (2014): Prediction the Groundwater Level of Bastam Plain (Iran) by Artificial Neural Network (ANN) and Adaptive Neuro-Fuzzy Inference System (ANFIS).– Water Resources Management, 28, 5433–5446. doi: 10.1007/s11269-014-0810-0

GAI, J. & HU, Y. (2018): Research on Fault Diagnosis Based on Singular Value Decomposition and Fuzzy Neural Network.– Shock and Vibration, ID 8218657. doi: 10.1155/2018/8218657

GONZALEZ CABRERA, J.M., FIDALGO MARTINEZ, M.R., MARTIN MATEOS, E.J. & VICENTE TAVERA, S. (2006): Study of the evolution of air pollution in Salamancam (Spain) along a five-year period (1994–1998) using HJ-Biplot simultaneous representation analysis.– Environmental Modelling & Software, 21, 61–68.

GULDAL, V. & TONGAL, H. (2009): Comparison of Recurrent Neural Network, Adaptive Neuro-Fuzzy Inference System and Stochastic Models in Egirdir Lake Level Forecasting.– Water Resources Management, 24, 105–128. doi: 10.1007/s11269-009-9439-9

GUZMAN, S.M., PAZ, J.O. & TAGERT, M.L.M. (2017): The Use of NARX Neural Networks to Forecast Daily Groundwater Levels.– Water Resources Management. 31, 1591–1603. doi: 10.1007/s11269-017-1598-5

HAO, Y., ZHANG, L., WEI, T. & WANG, X. (2010): Piecewise Analysis and Prediction of Spring Flow Based ON GM(1,1).– Advances in Grey Systems Research, doi: 10.1007/978-3-642-13938-3_14

JIANFEI, F., LUN, Z., ZHIGUO, L. & JIANYING, Y. (2012): Study on the Sustainable Utilization of Groundwater Resources in Hebei Plain.– Procedia Environmental Sciences, 12, 1071–1076. doi: 10.1016/j.proenv.2012.01.389

KANG, M.G. & MAENG, SJ. (2016): Gray Models for Real-Time Groundwater-Level Forecasting in Irrigated Paddy-Field Districts.– Journal of Irrigation and Drainage Engineering. doi: 10.1061/(ASCE)IR.1943-4774.0000940

Ratković et al.: Imputing missing data using grey system theory and the biplot method to forecast groundwater levels and yields 247

Geologia Croatica

KANG, M.S., KANG, M.G., PARK, S.W., LEE, J.J. & YOO, K.H. (2006): Application of grey model and artificial neural networks to flood forecasting.– Journal of the American Water Resources Association (JAWRA), 42, 473–486.

LIU, Q. & LUO, Y. (2010): The Non-equidistant Linear Weighting Grey Model NWGM (1,1) and Its Application.– International Journal of Simulation Systems, Science & Technology, doi: 10.5013/IJSSST.a.17.25.28

MAHMOD, W.E., WATANABE, K. & ZAHR-ELDEEN, A.A. (2013): Analysis of groundwater flow in arid areas with limited hydrogeological data using the Grey Model: a case study of the Nubian Sandstone, Kharga Oasis, Egypt.– Hydrogeology Journal, 21, 1021–1034. doi: 10.1007/s10040-013-0959-2

MALEŠEVIĆ, M. (1984): State Geological Map of SFRY, scale 1:100.000, Srbobran section (map and guide).– Federal Geological Survey, Belgrade.

MARTINS, W.O., PORSANI, M.J. & SILVA, M.G. (2016): Adaptive singular value decomposition filtering to enhance reflectors and geological structures in 3D seismic data.– Revista Brasileira de Geofisica, 34, 1–9. doi: 10.22564/rbgf.v34i2.797

MEHDIZADEH, S. (2020): Using AR, MA, and ARMA Time Series Models to Improve the Performance of MARS and KNN Approaches in Monthly Precipitation Modeling under Limited Climatic Data.– Water Resources Management (2020), 34, 263–282. doi: 10.1007/s11269-019-02442-1

MEIJAARD, J.P. (1993): Applications of the singular value decomposition in dynamics.- Computer Methods in Applied Mechanics and Engineering, 103, 161–173. doi: 10.1016/0045-7825(93)90044-X

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, 1315–1328. doi: 10.1007/s11269-014-0875-9

MOHAMMED, M., WATANABE, K. & TAKEUCHI, S. (2010): Grey model for prediction of pore pressure change.– Environmental Earth Sciences, 60, 1523–1534. doi: 10.1007/s12665-009-0287-y

MOHANTY, S., JHA, M.K., RAUL, S., PANDA, R. & SUDHEER, K. (2015): Using Artificial Neural Network Approach for Simultaneous Forecasting of Weekly Groundwater Levels at Multiple Sites.– Water Resources Management, 29, 5521–5532. doi: 10.1007/s11269-015-1132-6

NOURANI, V., MOGADDAM, A.A. & NADIRI, A.O. (2008): An ANN-based model for spatiotemporal groundwater level forecasting.– Hydrological Processes, 22, 5054–5066. doi: 10.1002/hyp.7129

POLOMČIĆ D., GLIGORIĆ Z., BAJIĆ D. & CVIJOVIĆ Č. (2017). A hybrid model for forecasting groundwater levels based on fuzzy C-mean clustering and singular spectrum analysis.– Water, 9/7, 541 (2017). doi: 10.3390/w9070541

POLOMČIĆ, D., BAJIĆ, D. & KRUNIĆ, O. (2014): Hydrodynamic analysis of the interaction of two operating groundwater source, case study: groundwater supply of Bečej.– Tehnika, 65, 597–603. doi: 10.5937/tehnika1404597P

POLOMČIĆ, D., KRUNIĆ, O. & RISTIĆ VAKANJAC, V. (2011): Hydrogeological and hydrodynamic characteristics of groundwater sources for the public water supply of Bečej (northern Serbia).– Annales Géologiques De La Péninsule Balkanique, 72, 143–157.

RAKHSHANDEHROO, G.R., VAGHEFI, M. & AGHBOLAGHI, M.A. (2012): Forecasting Groundwater Level in Shiraz Plain Using Artificial Neural Networks.– Arabian Journal for Science and Engineering, 37, 1871–1883. doi: 10.1007/s13369-012-0291-5

RISTIĆ VAKANJAC, V., ČOKORILO ILIĆ, M., PAPIĆ, P., POLOMČIĆ, D. & GOLUBOVIĆ, R. (2018): AR, CR and ARCR modeling for simulations and analyses of karst groundwater quality parameters.– Geoloski anali Balkanskog poluostrva, 79, 71–81. doi: 10.2298/GABP1879071R

TANG, Y.Q., CUI, Z.D., WANG, J.X., YAN, L.P. & YAN, X.X. (2008): Application of grey theory-based model to prediction of land subsidence due to engineering environment in Shanghai.– Environmental Geology, 55, 583–593. doi: 10.1007/s00254-007-1009-y

TERZIN, V. (1994): State Geological Map of SFRY, scale 1:100.000, Zrenjanin section (map and guide).– Federal Geological Survey, Belgrade.

WANG, C.N. & PHAN, V.T. (2016): An Improvement the Accuracy of Grey Forecasting Model for Cargo Throughput in International Commercial Ports of Kaohsiung.– Wseas transactions on business and economics, 11, 322–327. doi: 10.11648/j.ijber.20140301.11

WUNSCH, A., LIESCH, T. & BRODA, S. (2018): Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX).– Journal of Hydrology, 567, 743–758. doi: 10.1016/j.jhydrol.2018.01.045

XIAOJUN, D., JIANJUN, L. & HUAYING Z. (2015): Application of AR Model in the Analysis of Preearthquake Ionospheric Anomalies.– Mathematical Problems in Engineering, ID 157184. doi: 10.1155/2015/157184

YAN, Q. & MA, C. (2016): Application of integrated ARIMA and RBF network for groundwater level forecasting.– Environmental Earth Sciences, 75/5, 1–13. doi: 10.1007/s12665-015-5198-5

YONGPING, C., MIN, G., SHUNQI, P., HAIDONG, P., XIAN, Z. & ZHENGJIN, T. (2020): Application of auto-regressive (AR) analysis to improve short-term prediction of water levels in the Yangtze estuary.– Journal of Hydrology, 590, 125386. doi: 10.1016/j.jhydrol.2020.125386

YU, P.S., CHEN, C.J., CHEN, S.J. & LIN, S.C. (2001): Application of grey model toward runoff forecasting.– Journal of the American Water Resources Association (JAWRA), 37, 151–166. doi: 10.1111/j.1752-1688.2001.tb05482.x

ZHANG, L., LIU, Y. & ZHAO, F. (2018): Singular value decomposition analysis of spatial relationships between monthly weather and air pollution index in China.– Stochastic Environmental Research and Risk Assessment, 32, 733–748. doi: 10.1007/s00477-017-1491-z

ZHAO, B.B. & CHEN, Y.Q. (2010): Singular value decomposition (SVD) for extraction of gravity anomaly associated with gold mineralization in Tongshi gold field, Western Shandong Uplifted Block, Eastern China.– Nonlinear Processes Geophysics, 18, 103–109. doi: 10.5194/npg-18-103-2011

ZHAO-HUI, H. & JING, L. (2016): The Non-equidistance and Non-homogeneous Multivariable Grey NNMGM(1,n) Model and its Application.– Journal of Residuals Science & Technology, 13, 23.1–23.8.