Journal homepage: http://gwse.iheg.org.cn

Assessment of prediction performances of stochastic models: Monthly groundwater level prediction in Southern Italy

O Boulariah[1*], PA Mikhailov[2], A Longobardi[1], AN Elizariev[2], SG Aksenov[2]

[1] *Department of Civil Engineering*, *University of Salerno*, *Via Giovanni Paolo II*, 123, 84084, *Italy*.
[2] *Ufa State Aviation Technical University*, 450000, *Russian Federation*.

Abstract: Stochastic modelling of hydrological time series with insufficient length and data gaps is a serious challenge since these problems significantly affect the reliability of statistical models predicting and forecasting skills. In this paper, we proposed a method for searching the seasonal autoregressive integrated moving average (SARIMA) model parameters to predict the behavior of groundwater time series affected by the issues mentioned. Based on the analysis of statistical indices, 8 stations among 44 available within the Campania region (Italy) have been selected as the highest quality measurements. Different SARIMA models, with different autoregressive, moving average and differentiation orders had been used. By reviewing the criteria used to determine the consistency and goodness-of-fit of the model, it is revealed that the model with specific combination of parameters, SARIMA (0,1,3) $(0,1,2)_{12}$, has a high R^2 value, larger than 92%, for each of the 8 selected stations. The same model has also good performances for what concern the forecasting skills, with an average NSE of about 96%. Therefore, this study has the potential to provide a new horizon for the simulation and reconstruction of groundwater time series within the investigated area.

Keywords: Groundwater level forecast; Stochastic modelling; Southern Italy; Seasonality; Homogeneity

Received: 02 Nov 2020/ Accepted: 20 Apr 2021

Introduction

Groundwater evaluation has a profound impact on the management of water supplies, as groundwater is a major source of drinking water as well as irrigation water (Semiromi and Koch, 2019). The historical time series of groundwater level are the necessary tools for the study of groundwater drought features, to establish a relationship with the meteorological input and to potentially quantify the impact of climate conditions on the groundwater level by Longobardi and Van (2018). Depending on the temporal extension, lengths and frequency of data gaps of the time series, availability of neighbouring data at the regional scale and the characteristics of the study area, different techniques can be used to analyze groundwater levels. These techniques vary from numerical to mathematical, conceptual, and physical models, relying on the information used to explain the dynamic behaviour of aquifers and associated hydrological and hydrogeological variables. The Autoregressive Integrated Moving Average (ARIMA) family of models have been extensively used for the analysis of a wide range of water related issues (Longobardi and Villani, 2006; Faruk, 2010; Boulariah et al. 2019) including the groundwater level modelling (Ahn and Salas, 1997; Mirzavand and Ghazavi, 2014; Oikonomou et al. 2018). The complexity and heterogeneity of aquifer behaviour and parameters along with the poor input data requirements, makes the flexible statistical modelling a very attractive tool for the purpose (Suryanarayana et al. 2014; Oikonomou et al. 2018). Considering the detrending and the seasonality, the Box-Jenkins model is the most common approach to forecasting time series (Young, 1999; Adamowski, 2008). It has high accuracy in short-term forecasts (Takafuji et al.

*Corresponding author: O Boulariah, *E-mail address*: oboulariah@unisa.it

DOI: 10.19637/j.cnki.2305-7068.2021.02.008

Boulariah O, Mikhailov PA, Longobardi A, et al. 2021. Assessment of prediction performances of stochastic models: Monthly groundwater level prediction in Southern Italy. Journal of Groundwater Science and Engineering, 9(2): 161-170.

2019) and allows forecasting in the conditions of data shortage (Boulariah et al. 2019). In Kashan plain, Isfahan province of Iran, Mirzavand et al. (2015) used time series models and proved that the autoregressive moving average (ARMA) model can predict groundwater level change with high accuracy. In order to simulate the groundwater level in a coastal aquifer, Yang et al. (2017) explored three time series analysis approaches, i.e., Holt-Winters (HW), integrated time series (ITS) and seasonal autoregressive integrated moving average (SARIMA), showing that all three models can reliably predict the water table, but the SARIMA model showed highest reliability compared to the other two. Takafuji et al. (2019) compared geostatistical simulation (SGS) and the ARIMA model to forecast the depth of groundwater table in the Bauru Aquifer System (BAS), the results showed that the ARIMA model has better performance than the geostatistical model with higher accuracy and precision. In Italy, Nunno and Granata (2020) used the NARX neural network to predict groundwater level in Apulia region. They reported that very good results were obtained for the region of Apulia and suggested to use the NARX network in other areas characterized by Mediterranean climate and karst phenomena for groundwater prediction.

In this direction, the current paper aims to address the problem of predicting and forecasting groundwater table fluctuations in a specific area of the Mediterranean region, the Campania region, in southern Italy, characterized by temporally discontinuous observed time series, by using a statistical approach based on SARIMA tools.

1 Methodology

1.1 Time series homogeneity tests

Different statistical tests were used for homogeneity and trend analysis on time series data. Pettit's test and the student's test have been used in this study.

1.1.1 Pettitt's Test

In some cases, where the exact time of displacement is unknown, the use of the Pettitt's test allows us to determine the moment of change in the time series. So, the Pettit's test is based on the following rules:

If $X_1, X_2, X_3, \cdots X_n$ is a series of observable data that has a change point t, then $X_1, X_2, X_3, \cdots X_t$ has a distribution function $F_1(X)$, which differs from the distribution function of the second part of the series $X_{t+1}, X_{t+2}, X_{t+3}, \cdots X_H$. Nonparametric test statistics U_t can be written as follows:

$$U_t = \sum_{i=1}^{t}\sum_{j=t+1}^{n} sign(x_t - x_j) \quad (1)$$

Where: $sign(x_t - x_j) = \begin{bmatrix} 1, if\,(x_t - x_j) > 0 \\ 0, if\,(x_t - x_j) = 0 \\ -1, if\,(x_t - x_j) < 0 \end{bmatrix}$

Test statistics K and the corresponding confidence level ρ for the sample (n) can be described as:

$$K = Max|U_t| \quad (2)$$

When a certain value is more than the confidence level ρ

$$\rho = \exp\left(\frac{-K}{n^2 + n^3}\right) \quad (3)$$

The null hypothesis will be rejected. For the change point, the approximate probability of significance p is determined as follows:

$$p = 1 - \rho \quad (4)$$

Thus, in the place where the change point is located, the series is divided into two parts (Jaiswal et al. 2015).

1.1.2 Student's Test

The t-test refers to a parametric test in which the difference between the theoretical value and the average value is indicated by an alternative hypothesis, while the null hypothesis indicates that these values are equal to each other. This test has more statistical power than the nonparametric tests. t criterion is calculated by the formula (Cleophas et al. 2016):

$$t = \frac{\bar{x} - \mu}{s/\sqrt{n}} \quad (5)$$

Where: μ is the specified value; $\bar{x}$ is the mean of the samples; s is the standard deviation of the samples; n is the size of the samples.

These methods of testing the hypothesis on data homogeneity are selected on the basis of accuracy as well as the simplicity of utilization of results. Data analysis was performed by using the statistical software R studio.

1.1.3 Time Series prediction and forecasting models

The fundamental steps in forecasting and time series modeling, as described by (Valipour et al. 2012), are to plot the time series and determine its basic characteristics. Regarding various features such as trend and seasonality, it is important to look for potential signs that the time series has changed over the investigation period. Next, trend or seasonality should be eliminated by differentiating or by fitting model. The main goal of this process is to produce a set of stationary residuals (Valipour et al. 2013). As a final step, a forecasting model for the residuals should be developed. Occasionally, several potential models can be found, and in this case, an additional analysis must be performed to determine the best model for deployment.

Sometimes, potential models can be eliminated based on how well they fit over historical data. It is possible that a poorly adapted model may produce good forecasts. Montgomery et al. (2015) has shown that it is necessary to validate the performance of a model from the previous step. This will likely involve some type of split sampling or cross-validation procedure. The objective of this step is to select a suitable forecast model.

In addition, it is interesting to know the differences between the original data and the forecasted values by the original scale model. The methodology for determining the optimal seasonal ARIMA model consists of six consecutive steps as described above. Fig. 1 demonstrates the step-by-step process to forecast time series from data analysis to the estimation of the corresponding model.

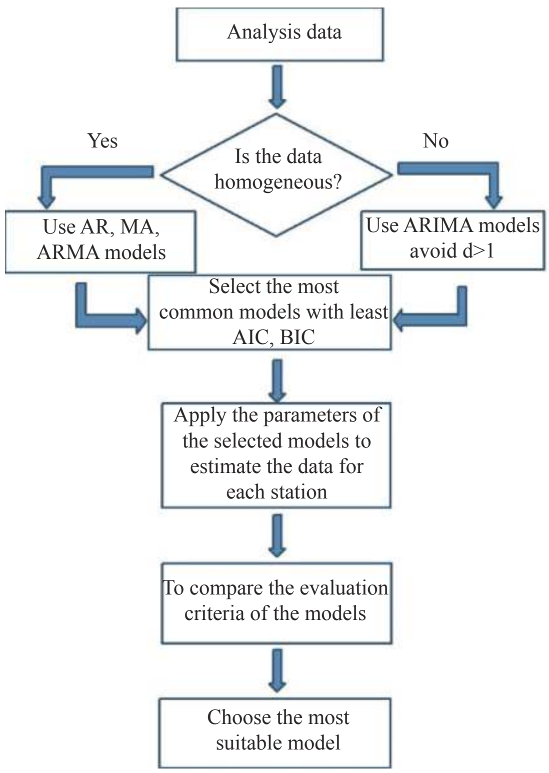

Fig. 1 Flowchart of applied methodology for forecasting time series

The use of ARMA models has proven itself in the forecasting and analysis of hydrological time series, which is confirmed by Cleophas and Zwinderman (2016), Balasmeh et al. (2019) and Mombeni et al. (2013). As mentioned above, the ARMA model is a combination of the autoregressive (AR) model and the moving average (MA) model (Kumar and Rathnam, 2019). AR part indicates that the variable of interest regresses according to its previous values. MA part shows that a linear combination of a limited number of residuals is a regression error. In other words, the deviation of a dependent variable from its average value is a linear combination of current and past values of the random perturbation vector. Thus, the ARMA model allows one to make a forecast that depends on both the current and past values of variable as well as the magnitude of normal disturbances of the current and past values. The stationarity of data usage is an essential criterion for the use of such models.

The chosen model is based on manually selected parameters that depends on the values of ACF and PACF graphs, which minimizes the Akaike and Swartz criteria, and is also based on essential coefficients. The model will not be accepted if the resulting model does not meet certain mathematical conditions (Valipour et al. 2013a). The forecasted values have a predictive interval that increases as the forecast horizon increases. The forecast intervals are based on the theory that there is no autocorrelation of the residues and their normal distribution. If one of the conditions is not satisfied, the forecast may turn out to be incorrect. Therefore, A graph and residual histogram are built before the forecast, and together with the Lung-Box test, a visual analysis that allows us to draw conclusions whether the model fits or not.

The Lung-Box test allows us to check the hypothesis that the autocorrelation coefficient differs from zero. The null hypothesis of the test indicates that the remnants of the model are white noise. An alternative hypothesis suggests that the data is not random. The formula for calculating Q-statistics is as follows :

$$Q^{*} = T(T+2)\sum_{k=1}^{m}\frac{r_k^2}{T-k} \tag{6}$$

Where: T is the sample size; m is the maximum lag length; r_k is the sample autocorrelation coefficient for lag k;

The critical value is given as follows:

$$X_{\kappa p}^2 = X_{1-\alpha,m}^2 \tag{7}$$

Where: $X_{1-\alpha,m}^2$ is quantile of chi-square distribution with m degrees of freedom. If $Q^{*} > X_{\kappa p}^2$, the null hypothesis will be rejected; if $Q^{*} < X_{\kappa p}^2$, the null hypothesis will be accepted at a given level of significance.

The Lung-Box test is based on the Box-Pierce statistics. Therefore, it has the same asymptotic distribution and gives comparatively similar results for relatively large numbers. The criterion does not lose its consistency even if the process does not have a normal distribution. When the data under study is non-stationary, the box-Jenkins models are used. To reduce the data to a stationary form, a method of taking consecutive differences is used. If the series becomes stationary after these transformations, we can use the ARIMA model to forecast. In this case, the initial step of differentiation can be applied one or several times to eliminate the unsteadiness (Valipour et al. 2013b). There are four consecutive stages to determine the Box- Jenkins model:

(1) Model identification-the process of selecting a model that is best suited to the process in question.

(2) Model assessment-the use of regression methods to obtain the parameters included in the model.

(3) Model test-check the adequacy of the model using tests for normality and autocorrelation of residues.

(4) Model application-use of the model focus forecasting. (The design of the technique is presented in Fig. 2).

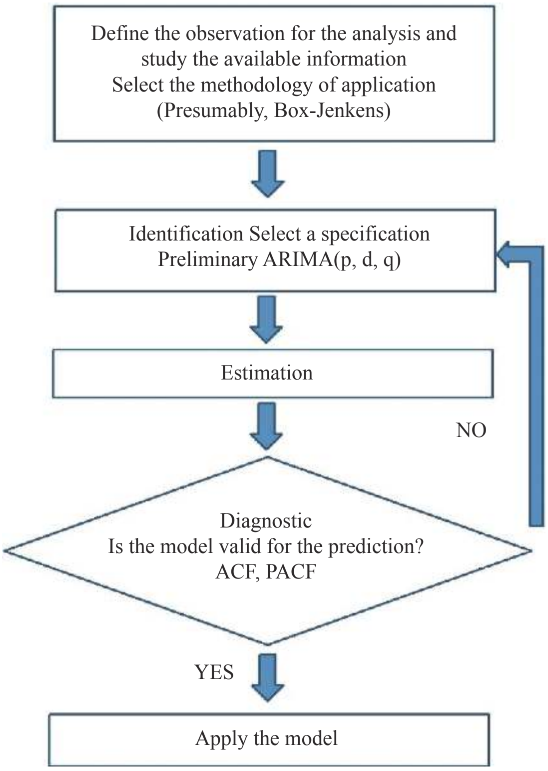

Fig. 2 ARIMA model prediction method (Chatfield et al.1973)

In the case where the data contain a seasonal component, the SARIMA model is used for forecasting. The difference between SARIMA and ARIMA model is to add three parameters (P, D, Q) to predict data seasonality. The equation of the ARIMA/SARIMA model in general looks like this:

$$\varnothing_p(B^s)\varphi_P(B)(1-B)^d(1-B^s)^D Z_t = \theta_q(B)\vartheta_Q(B^s)e_t \quad (8)$$

$$\varnothing_p(B^s) = 1-\varnothing_1 B^s-\varnothing_2 B^{2s}-\cdots-\varnothing_p(B^{ps}) \quad (9)$$

$$AR(p)-\varphi_p(B) = 1-\varphi_1 B-\varphi_2 B^2-\cdots-\varphi_P B^P \quad (10)$$

$$\mathrm{MA}(q)-\theta_q(B) = 1-\theta_1 B-\theta_2 B^2-\cdots-\theta_q B^q \quad (11)$$

Where: AR(p) is the operator that converges to provide the operator (MA) with stationary conditions, which is a polynomial of the order q in B, such that each value is the sum of the previous white noise values of q+1.

MA (q) is the operator which converges to ensure the reversibility

$$\vartheta_Q(B^s) = 1-\vartheta_1 B^s-\theta_2 B^{2s}-\cdots-\vartheta_Q B^{Qs} \quad (12)$$

1.1.4 Goodness of fit

In order to analyze the performance of these models, various efficiency indices, such as Nash-Sutcliffe Efficiency (NSE) (Nash and Sutcliffe, 1970), R^2 determination coefficient, root mean square error (RMSE) (Singh et al. 2005), Mean Absolute Error (MAE) (Willmott et al. 2005), Index of Agreement (Willmott, 1981), Mean Squared Derivative Error (Willmott et al. 2005) and the percent bias (%) (Singh et al. 2005). In addition, Akaike Information Criterion (AIC) and Bayesian Information Criterion (BIC) were used to evaluate the maximum likelihood of models. Moriasi et al. (2007) suggested a rating method for this purpose to classify models into four types: very good, good, satisfactory and unsatisfactory.

2 Study area and datasets

The Campania region in southern Italy is situated between 40.0°-41.5° N and 13.5°-16.0° E, extending from the Apennine Mountains to the Mediterranean Sea with steadily decreasing elevations from the interior to the coastline, covering approximately 14 000 km^2 (Fig. 3, right panel). The climate of study area is usually seasonal, with some apparent variations depending on the region. With the greater amount of precipitation observed during the winter months, the seasonality is very marked. The mean annual rainfall ranges from 600 mm to 2 400 mm in the study region (Fig. 3, left panel), while the average

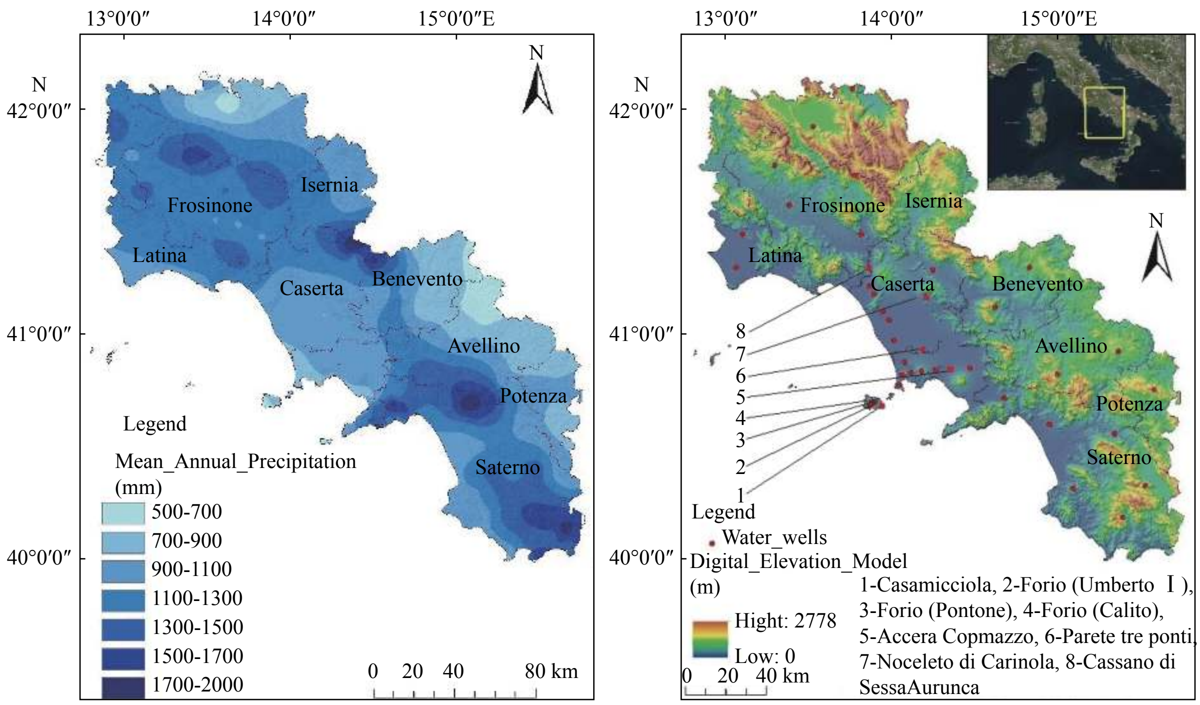

Fig. 3 Digital elevation model with location of water wells (right panel) and mean annual precipitation (left panel) in the study area

annual temperature is about 17 °C.

Available data consist of monthly groundwater level time series for over 44 sites located across the Campania region and the Lazio region, the data were reviewed and analyzed in this paper from 1926 to 1994. The water wells database was provided by the national hydrographic and mareograph service SIMN. After a data quality control performed using different statistical tests, only 8 out of these stations were effectively used in the area under investigation.

3 Results and discussion

To determine the parameters of the model, the auto ARIMA function of the R software was used, the function allows us to automatically select the model with the lowest AIC, BIC indices, which significantly reduces the time for the selection of parameters and model order. When evaluating the observed values, a restriction in the enumeration of parameters was used. Models with a step of differentiation (d>1, D>1) were excluded. This restriction is based on studies conducted by Hamilton (1994) and Mohammadi et al. (2005), which recommended that the differentiation order for time series should not be more than 1, otherwise the results will be far from reality. The following is an example of estimating the observed values for data from a measuring station in Casamicciola. We applied the SARIMA model (1,1,2) $(0,1,1)_{12}$, with the lowest AIC and BIC (AIC = −1 226.74 and BIC = −1 207.5). A comparison of the predicted and observed values is presented in Fig. 4.

As it can be seen from the picture (Fig. 4), the data have an increasing trend, which is confirmed by the tests for stationarity. According to the Pettitt test, the series is heterogeneous and a shift is

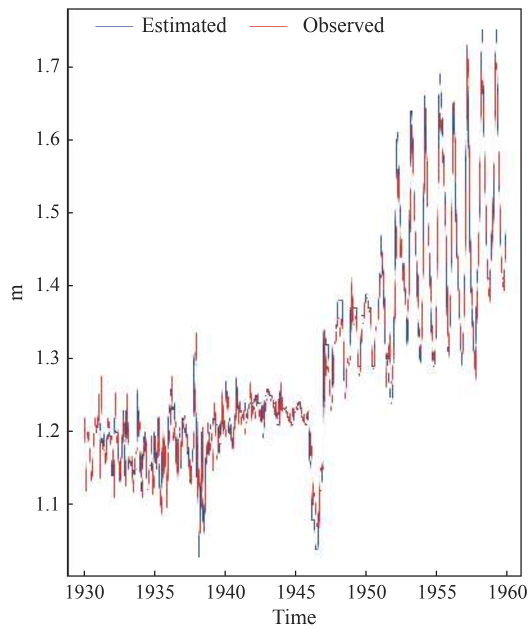

Fig. 4 Predicted and observed values of the ground water level

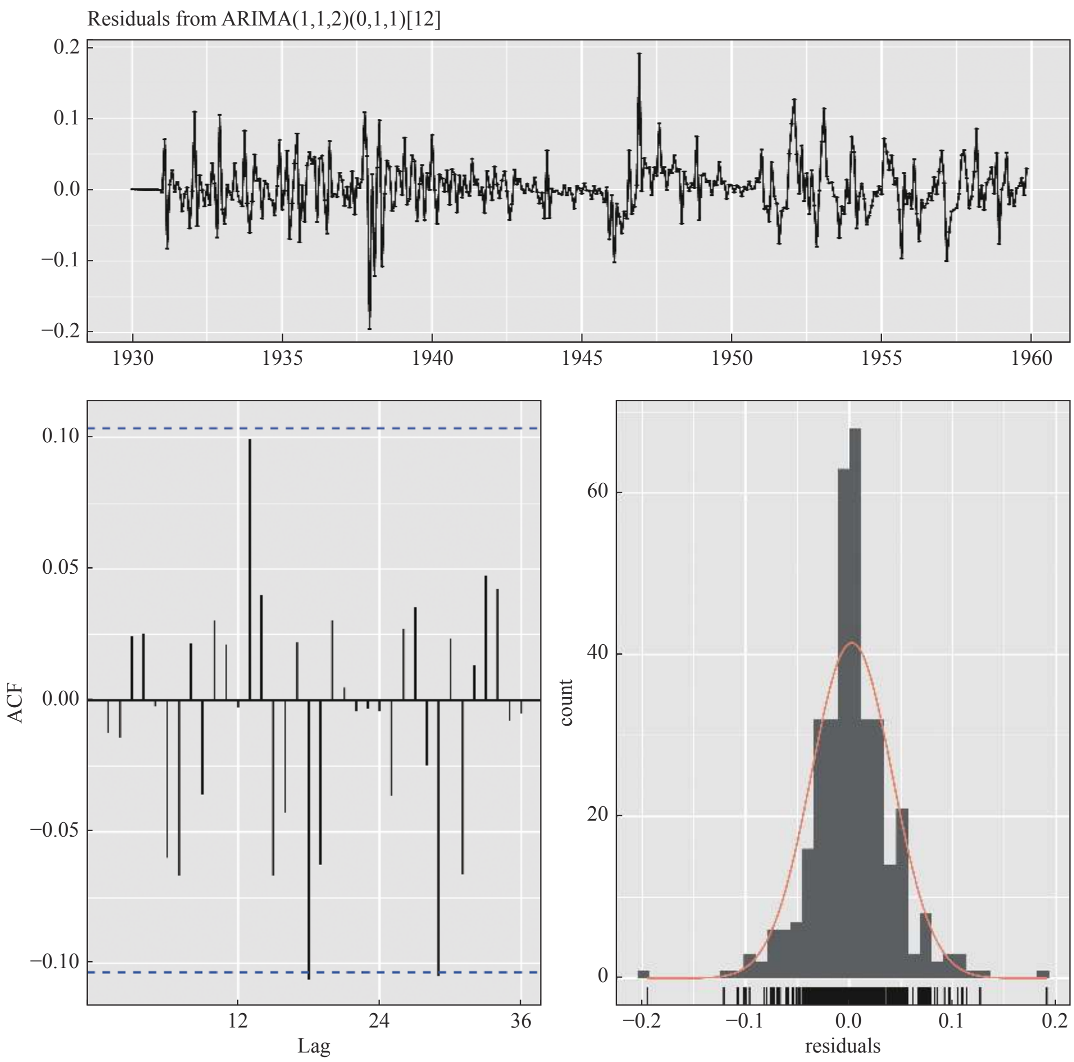

Fig. 5 Autocorrelation function graph for the residues of the SARIMA model (1,1,2) $(0,1,1)_{12}$

observed in August 1946. The model describes the data well enough, which is confirmed by the efficiency indices between the estimated and observed values. A high correlation value between graphs indicates that the prediction is highly accurate. The adequacy of the model was verified using the residual analysis shown in Fig. 5.

Residuals look like white noise and correspond to a normal distribution and the values of the autocorrelation lags are within the confidence interval, which indicates the absence of autocorrelation in the residues. The histogram of residuals shows symmetric distribution relative to zero, which indicates the absence of systematic errors. Those results suggest that this model is well suited for evaluating the data in question and can be used for prediction and forecasting. In a similar way, models were found for predicting all 8 stations, the results are presented in Table 1.

The analysis of the resulting models allowed us to determine the three most frequent forecasting models, which are ARIMA (0,1,3) $(0,1,2)_{12}$, ARIMA (1,1,2) $(0,1,1)_{12}$, ARIMA (1,1,1) $(0,1,1)_{12}$. The three recurring models were applied for each of the stations. The model parameters were calibrated using various statistical indices of model quality, such as (NSE) (Nash and Sutcliffe, 1970),

Table 1 Forecasting models Results based on AIC and BIC criteria

Name of station	Parameters of model	AIC	BIC
Acera Capomazzo	$(1,1,2)(0,1,2)_{12}$	514.44	542.39
Casamicciola	$(1,1,2)(0,1,1)_{12}$	−1 228.75	−1 209.5
Cassano di Sessa Aurunca	$(0,1,3)(0,1,2)_{12}$	914.41	942.54
Forio(Calitto)	$(1,1,2)(0,1,1)_{12}$	607.55	630.75
Forio(Pontone)	$(0,1,3)(0,1,2)_{12}$	−571.39	−547.92
Forio(Umberto I)	$(1,1,1)(0,1,1)_{12}$	−39.8	−24.01
Nocelleto di Carinola	$(1,1,1)(0,1,1)_{12}$	−377.44	−309.26
Parete(tre ponti)	$(1,1,2)(0,1,1)_{12}$	−702.58	−682.21

correlation coefficient R^2, root mean square error (RMSE) (Singh et al. 2005) and percentage bias (%), etc. The goodness-of-fit assessment of the models was carried out on the basis of 11 statistical criteria presented below (Table 2).

The analysis of Table 2 shows high values of Pearson correlation for each model and the average value is 0.94. Such a high correlation shows that each of the models describes the data well enough to ensure high accuracy of prediction.

However, the SARIMA model (0,1,3) $(0,1,2)_{12}$ is the best for predicting the groundwater level for each of the 8 stations in the Campania region, since it has the highest correlation between the observed and forecasted values and the lowest errors than two other models. Using this model to forecast

Table 2 The comparison of different SARIMA models based on accuracy measures

	Parete (tre ponti)	Nocelleto di carinola	Forio (Umberto I)	Forio (Pontone)	Forio (Calitto)	Cassano di Sessa Aurunca	Casamicci ola	Accera Copmazzo
				SARIMA $(0,1,3)(0,1,2)_{12}$				
NSE	0.944	0.996	0.920	0.949	0.985	0.999	0.932	0.984
MAE	0.065	0.110	0.131	0.067	0.217	0.255	0.026	0.214
RMSE	0.107	0.183	0.218	0.108	0.363	0.420	0.037	0.331
Pearson cor.	0.971	0.935	0.960	0.974	0.962	0.930	0.965	0.983
MSE	0.011	0.032	0.047	0.011	0.121	0.176	0,001	0.109
d	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
BIAS$	−0.026	0.027	−0.006	−0.012	−0.046	−0.045	0.036	0.003
MSDE	4.7E-04	4.3E-06	1.5E-05	3.1E-09	0.006	0.101	2.2E-06	0.016
R^2	0.944	0.874	0.922	0.949	0.925 7	0.865	0.932	0.967
AIC	−678.64	−321.22	−36.87	−571.61	625.55	914.41	−1 203.63	532.08
BIC	−654.2	−294.18	−13.18	−547.92	652.85	941.39	−1 180.55	559.97
				SARIMA $(1,1,2)(0,1,1)_{12}$				
NSE	0.943	0.996	0.920	0.947	0.689	0.999	0.930	0.984
MAE	0.064	0.109	0.129	0.068	0.529	0.255	0.026	0.211
RMSE	0.108	0.182	0.218	0.110	1 625 774	0.422	0.037	0.329
Pearson cor.	0.971	0.934	0.960	0.973	0.553	0.929	0.964	0.983
MSE	0.012	0.032	0.047	0.012	0.419	0.178	0.001	0.107
d	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
BIAS$	−0.018	0.032	−0.008 8	−0.013	−2.340	−0.046	−0.117	0.057
MSDE	0.000 1	2.6E-06	1.3E-05	3.1E-09	0.198	0.106	2.2E-06	0017
R^2	0.943	0.873	0.921	0.948	0.306	0.863	0.931	0.967
AIC	−702.58	−327.03	−37.95	−562.72	607.55	920.63	−1 226.74	519.45
BIC	−682.21	−304.5	−18.2	−542.97	630.3	943.12	−1 207.5	542.69
				SARIMA $(1,1,2)(0,1,1)_{12}$				
NSE	0.943	0.996	0.920	0.946	0.985	0.999	0.940	0.984
MAE	0.063	0.110	0.129	0.069	0.219	0.257	0.024	0.212
RMSE	0.108	0.183	0.218	0.111	0.365	0.424	0.034	0.330
Pearson cor.	0.971	0.934	0.960	0.973	0.962	0.928	0.970	0.983
MSE	0,011	0.032	0.047	0.012	0.122	0.180	0.001	0.107
d	0.999	0.999	0.999	0.999	0.999	0.999	0.999	0.999
BIAS$	−0.018	0.033	−0.008	−0.013	−0.042	−0.047	0.048	0.061
MSDE	0,0001	2.4E-06	1.3E-05	3.1E-09	0.005	0.109	2.2E-06	0017
R^2	0.943	0.873	0.921	0.947	0.862	0.862	0.941	0.967
AIC	−672.01	−327.44	−39.8	−561.6	629.36	924.5	−1 194.32	518.6
BIC	−655.71	−309.41	−24	−545.8	647.56	942.491	−1 178.931	537.19

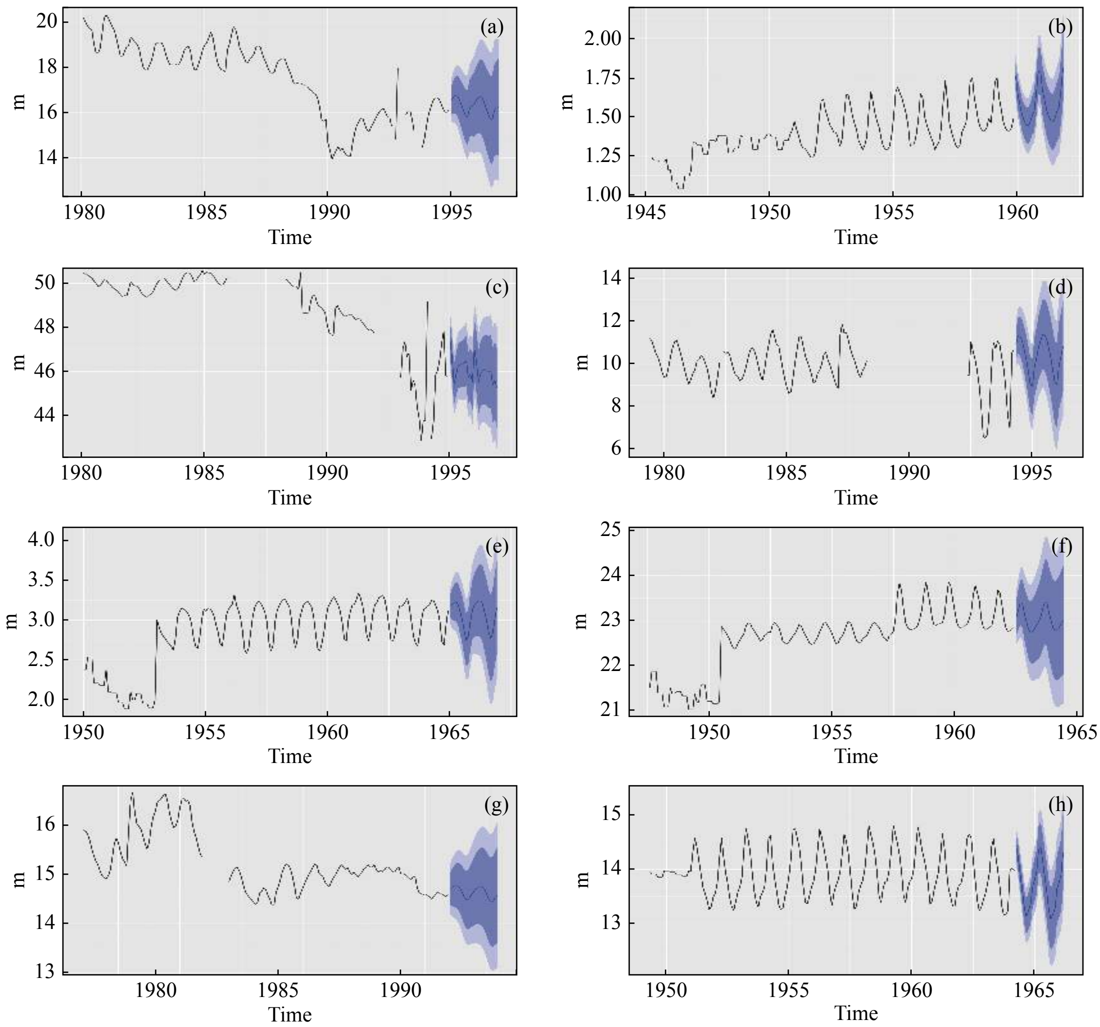

Fig. 6 Forecasting using SARIMA(0,1,3)(0,1,2)$_{12}$; (a) Acera Capomazzo, (b) Casamicciola, (c) Cassano di Sessa Aurunca, (d) Forio(Calitto), (e) Forio(Pontone), (f) Forio(Umberto I), (g) Nocelleto di Carinola, (h) Parete(tre ponti) scrivi cosa sono le fasce colorate

groundwater level behavior for 24 months, the obtained results are shown in Fig. 6, where blue and gray colors highlight the 80% and 95% confidence intervals respectively.

As shown in Fig. 6, this model is very suitable for predicting time series in each region; data is well predicted despite the presence of trend and seasonality. For clarity, and according to classification by Moriasi et al. (2007), Table 3 shows the values of some statistical criteria and their average values when using the SARIMA model (0,1,3) (0,1,2)$_{12}$.

The average value of the indices for the selected

Table 3 Statistical index values for the selected SARIMA model (0,1,3) (0,1,2)$_{12}$

	NSE	BIAS%	R^2	*d*	r
Accera Copmazzo	0.98	0.003	0.96	0.99	0.98
Casamicciola	0.93	0.03	0.93	0.99	0.96
Cassano di Sessa Aurunca	0.99	−0.04	0.86	0.99	0.93
Forio Calitto	0.98	−0.04	0.92	0.99	0.96
Forio Pontone	0.94	−0.012	0.94	0.99	0.97
Forio Umberto I	0.92	−0.006	0.92	0.99	0.96
Nocelleto di carinola	0.99	0.02	0.87	0.99	0.93
Parete tre ponti	0.94	−0.02	0.94	0.99	0.97
Model Quality (Very good)	**0.75< NSE<1.00**	**PBIAS<±10**	**$0.75 < R^2 \leq 1.0$**	**1**	**r > 0.7**

model shows that this model describes the data accurately and allows us to predict and forecast the selected time series with high accuracy and precision.

4 Conclusions

Ground water level forecasting is an integral tool in the management, implementation, and operation of drinking water supply systems. The use of the Box-Jenkins technique is a reasonably accurate approach for groundwater time series modeling. The results show that SARIMA models predict monthly groundwater levels with high accuracy. The seasonal autoregressive integrated moving average models are sensitive to extreme values and data periodicity. The procedure used to carry out the work simplifies the search for the desired combination of prediction parameters. In fact, the most suitable parameters for predicting the groundwater level behavior of the Campania region is the (0,1,3) $(0,1,2)_{12}$, which is confirmed by the analysis of statistical indices. The SARIMA models obtained have furthermore a sufficiently high forecasting accuracy although there is basically no way to improve them, since they select a significant part of the information from the data. The relevance of this model for predicting the groundwater level in each of the 8 selected stations proves the possibility of using the model to simulate the situation in the entire region using the data of remaining stations.

References

Adamowski J. 2008. Development of a short-term river flood forecasting method for snowmelt driven floods based on wavelet and cross-wavelet analysis. Journal of Hydrology, 353: 247-266.

Ahn H, Salas J. 1997. Groundwater head sampling based on stochastic analysis. Water Resources Research, 33: 2769-2780.

Balasmeh A, Babbar O, Karmaker T. 2019. Trend analysis and ARIMA modeling for forecasting precipitation pattern in Wadi Shueib catchment area in Jordan. Arabian Journal of Geosciences, 12: 27-27.

Boulariah O, Meddi M, Longobardi A. 2019. Assessment of prediction performances of stochastic and conceptual hydrological models: Monthly stream flow prediction in north-western Algeria. Arabian Journal of Geosciences, 12: 792-792.

Chatfield C, Prothero DL. 1973. Box - Jenkins seasonal forecasting: Problems in a case - study. Journal of the Royal Statistical Society: Series A (General), 136(3): 295-315.

Cleophas T, Zwinderman A. 2016. One-Sample Continuous Data (One-Sample T-Test, One-Sample Wilcoxon Signed Rank Test, 10 Patients). SPSS for Starters and 2nd Levelers: 3-6.

Faruk DÖ. 2010. A hybrid neural network and ARIMA model for water quality time series prediction. Engineering applications of artificial intelligence, 23(4), 586-594.

Hamilton J. 1994. Time series analysis. Princeton University Press. 1-820.

Jaiswal R, Lohani A, Tiwari H. 2015. Statistical Analysis for Change Detection and Trend Assessment and Climatological Parameters. Environmental Processes: 729-749.

Kumar KS, Rathnam EV. 2019. Analysis and prediction of groundwater level trends using four variations of Mann Kendall tests and ARIMA modelling. Journal of the Geological Society of India, 94: 281-289.

Longobardi A, Villani P. 2006. Seasonal response function for daily stream flow investigation. Physics and Chemistry of the Earth, 31: 1107-1117.

Longobardi A, Van Loon A. 2018. Base flow index vulnerability to variation in dry spell length for a range of catchment and climate properties. Hydrological Processes, 32: 2496-2509.

Montgomery DC, Jennings CL, Kulahci M. 2015. Introduction to time series analysis and forecasting. John Wiley & Sons.

Mirzavand M, Ghazavi R. 2014. A stochastic modelling technique for groundwater level forecasting in an arid environment using time series methods. Water resources management, 29(4): 1315-1328.

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.

Mohammadi K, Eslami H, Dayani DS. 2005. Comparison of regression, ARIMA and ANN models for reservoir inflow forecasting using snowmelt equivalent (a case study of Karaj). Journal of Agricultural Science and Technology, 7: 17-30.

Mombeni H, Rezaei S, Nadarajah S, et al. 2013. Estimation of water demand in Iran based on SARIMA models. Environmental Modeling & Assessment, 18: 559-565.

Moriasi D, ARnoid G, Kuijk M, et al. 2007. Model evaluation guidelines for systematic quantification of accuracy in watershed simulations. Transactions of the ASABE, 50: 885-900.

Nash J, Sutcliffe J. 1970. River flow forecasting through conceptual models part I — A discussion of principles. Journal of Hydrology, 10: 282-290.

Nunno FD, Granata F. 2020. Groundwater level prediction in Apulia region (Southern Italy) using NARX neural network. Environmental Research, 190: 110062.

Oikonomou PD, Alzraiee AH, Karavitis CA, et al. 2018. A novel framework for filling data gaps in groundwater level observations. Advances in Water Resources, 119: 111-124.

Semiromi M, Koch M. 2019. Reconstruction of groundwater levels to impute missing values using singular and multichannel spectrum analysis: application to the Ardabil Plain, Iran. Hydrological sciences journal, 64: 1711-1726.

Singh J, Knapp H, Arnold J, et al. 2005. Hydrologic modeling of the Iroquois River watershed using HSPF and SWAT. J. American Water Resources Assoc, 41(2): 361-375.

Suryanarayana C, Sudheer C, Mahammood V, et al. 2014. An integrated wavelet-support vector machine for groundwater level prediction in Visakhapatnam, India. Neurocomputing, 145: 324-335.

Takafuji E, Rocha M, Manzione R. 2019. Groundwater level prediction forecasting and assessment of uncertainty using SGS and ARIMA Models: A case study in the Bauru Aquifer System (Brazil). Natural Resources Research, 28: 487-503.

Valipour M, Banihabib M, Behbahani S. 2013. Comparison of the ARMA, ARIMA, and the autoregressive artificial neural network models in forecasting the monthly inflow of Dez dam reservoir. Journal of Hydrology, 476: 433-441.

Valipour M. 2013a. Increasing irrigation efficiency by management strategies: cutback and surge irrigation. ARPN Journal of Agricultural and Biological Science. 8(1):35-43.

Valipour M. 2013b. Necessity of irrigated and rainfed agriculture in the world. Irrigation & Drainage Systems Engineering. S9, e001.

Valipour M. 2012. Ability of Box-Jenkins models to estimate of reference potential evapotranspiration (A case study: Mehrabad Synoptic Station, Tehran, Iran). IOSR Journal of Agriculture and Veterinary Science, 1.5: 1-11.

Willmott C. 1981. On the validation of models. Physical Geography, 2: 184-194.

Willmott CJ, Matsuura K. 2005. Advantages of the mean absolute error (MAE) overthe root mean square error (RMSE) in assessingaverage model performance, 30, 79-82.

Young P. 1999. Nonstationary time series analysis and forecasting. Progress in Environmental Science, 1: 3-48.

Yang Q, Wang Y, Zhang J, et al. 2017. A comparative study of shallow groundwater level simulation with three time series models in a coastal aquifer of South China. Applied Sciences, 7: 689-698.