CrossMark

Groundwater Level Prediction/Forecasting and Assessment of Uncertainty Using SGS and ARIMA Models: A Case Study in the Bauru Aquifer System (Brazil)

Eduardo Henrique de Moraes Takafuji [1,3] **Marcelo Monteiro da Rocha** [1]
and Rodrigo Lilla Manzione [2]

Best water management practices should involve the prediction of the availability of groundwater resources. To predict/forecast and consequently manage these water resources, two known methods are discussed: a time series method using the autoregressive integrated moving average (ARIMA) and a geostatistical method using sequential Gaussian simulation (SGS). This study was conducted in the Ecological Station of Santa Barbara (EEcSB), located at the Bauru Aquifer System domain, a substantial water source for the countryside of São Paulo State, Brazil. The relevance of this study lies in the fact that the 2013/2014 hydrological year was one of the driest periods ever recorded in São Paulo State, which was directly reflected in the groundwater table level behavior. A hydroclimatological network comprising 49 wells was set up to monitor the groundwater table depths at EEcSB to capture this response. The traditional time series has the advantage that it has been created to forecast and the disadvantage that an interpolation method must also be used to generate a spatially distributed map. On the other hand, a geostatistical approach can generate a map directly. To properly compare the results, both methods were used to predict/forecast the groundwater table levels at the next four measured times at the wells' locations. The errors show that SGS achieves a slightly higher level of accuracy and considered anomalous events (e.g., severe drought). Meanwhile, the ARIMA models are considered better for monitoring the aquifer because they achieved the same accuracy level as SGS in the 2-month forecast and a higher precision at all periods and can be optimized automatically by using the Akaike information criterion.

KEY WORDS: Geostatistics, Time series, Groundwater monitoring, Bauru aquifer system, Ecological Station of Santa Barbara (Brazil).

INTRODUCTION

Currently, large volumes of data or information about certain phenomena are being collected over a long period of time and over a wide spatial distribution, such as atmospheric information retrieved at meteorological or air quality stations, health data at hospitals or health posts, and groundwater quality and levels at water wells. Many kinds of data and/or

[1] Institute of Geosciences - University of São Paulo (IGc-USP), Rua do Lago, 562 Cidade Universitária, São Paulo, SP 05508-080, Brazil.
[2] School of Sciences and Engineering - São Paulo State University (FCE-UNESP), Rua Domingos da Costa Lopes, 780 Jd. Itaipu, Tupã, SP 17602-496, Brazil.
[3] To whom correspondence should be addressed; e-mail: eduardo.takafuji@usp.br

information are, nowadays, being captured by remote sensing from satellites, geospatial networks, and ground sensors, among others. With huge volumes of data and/or information available, choosing the best method to analyze them is not a straightforward issue even for experienced modelers. Theories and applications of probabilistic or stochastic hydrological approaches have been developed over the last 30 years; however, their applications to real problems have been limited and they have not become routine tools in hydrological modeling (Dagan 2002; Renard 2007; Zhang and Zhang 2004). Future scenarios involving climate uncertainty have also boosted the popularization of these approaches, especially for the interpretation of results from stochastic experiments. The resistance of certain sectors of the academia to accept different equiprobable results is justified in areas of knowledge where precision is a preponderant criterion for acceptability. In environmental analyzes, however, where there is a need for tools capable of describing and predicting complex environmental processes, the study of environmental time series is fundamental to the larger goal of sustainability and adaptation (Cressie and Holan 2011).

Traditional time series analysis is the first method to consider when modeling chronological data. However, other approaches, such as geostatistical methods, which have been widely developed and applied in the last decades, can contribute to enhancing the results of forecasting experiments. Kyriakidis and Journel (1999) explained that space–time data are traditionally analyzed through models that were initially developed for spatial or temporal distributions. The joint space–time dependence is often neither fully modeled nor exploited in the estimation or forecasting of unknown values at unsampled locations. Moreover, modeling of spatiotemporal distributions resulting from dynamic environmental processes evolving in both space and time is critical in many scientific and engineering fields. On the other hand, geostatistical spatiotemporal models are becoming more popular because they provide probabilistic framework for data analysis, as well as, predictions that build on the joint spatial and temporal dependence between observations. In contrast, for space–time data, the complexity of geostatistics has discouraged its applications.

Despite the complexity of spatiotemporal geostatistics framework, this paper compares two well-known methods due to their practical applications,

in order to make it possible to implement them in monitoring groundwater levels and, if necessary, improve the spatiotemporal geostatistics approach. The first method is an extrapolation method (time series) and the second is an interpolation method (geostatistical simulation), both of which are used to forecast groundwater table level. Initially, brief reviews of the time series and geostatistical simulation are presented. The time series method reproduces the trend, seasonal, cyclical and irregular components of the original data, whereas the sequential Gaussian simulation (SGS) reproduces the global spatial variance, and its simulations can be used as long as the variogram is representative. Because the SGS is sequential, it uses previously simulated data to simulate the future. The cost of the long-term extrapolation is that the SGS tends to honor the global mean and variance. Thus, the seasonality cannot be reproduced properly after a few steps. The case study presented here is based on the comparison of SGS and autoregressive integrated moving average (ARIMA) procedures on four ''future'' dates (October 16, October 29h, November 13, and December 3) at 49 water wells located at the Ecological Station of Santa Barbara. The study addresses the problem of calculating future groundwater level (depth) and determining the degree of uncertainty of this calculation. The first approach is to run SGS on the four future dates at the specific 49 water well locations. The second approach is to use time series forecasting with ARIMA models for the same geographical points. Both methods calculate the expected values and the confidence level; this study considers 10% (P(10)) and 90% (P(90)) scenarios. Finally, both methods are compared.

APPROACHES FOR FORECASTING GROUNDWATER TABLE DEPTHS

There are several applications of time series analysis in hydrogeology. Yi and Lee (2004) filled and completed irregular hydraulic head series from precipitation data. Hatch et al. (2006) evaluated the interaction between groundwater and surface water, whereas Von Asmuth et al. (2008) evaluated the effect of multiple stresses such as baseline flow, pumping, vegetation, climate, and dams on the behavior of groundwater levels. More complex time series models can estimate groundwater recharge (Yihdego and Webb 2011), capture nonlinear soil drainage behavior (Peterson and Western 2014), or

even determine recharge response time from precipitation (Hocking and Kelly 2016; Manzione et al. 2017). Moreover, Mackay et al. (2015) used seasonal rainfall forecasts and a lumped groundwater model to simulate groundwater levels up to 3 months into the future. Furthermore, geostatistical methods have been used as data-driven variable map generators, usually combined with a forecast method. Bierkens (2001) applied stochastic differential equations (SDE) and kriging with a digital elevation model (DEM) as external drift to predict soil saturation. Castrignanò and Buttafuoco (2004) applied SGS to assess the probability that water content does not exceed a critical threshold that might cause water stress to a forest of Laricio pine (Italy). Lee et al. (2007) applied SGS and transition probability indicator simulation to create conditional realizations of alluvial fan aquifer systems in the Lawrence Livermore National Laboratory area to simulate a pumping test. Omran (2016) also used SGS to simulate early predictions of groundwater table level fluctuations in North Sinai, Egypt. Singh and Katpatal (2017) proposed a method to design an optimized groundwater level monitoring network using a multi-criteria analysis (weighted overlay, analytical hierarchical process, and fuzzy) and a geostatistical method (ordinary kriging).

Comparable methods have been proposed by many authors in order to analyze aquifers spatiotemporally; some of these methods are briefly reviewed here. Knotters and Bierkens (2001) and Bierkens et al. (2001) implemented the regionalized autoregressive exogenous variable (RARX) model, which uses auxiliary information such as soil profile, topographic maps, and DEMs. The Predefined Impulse Response Function In Continuous Time (PIRFICT) model was applied by Manzione et al. (2008) to map groundwater table depths and quantify the associated uncertainty, in order to indicate the extreme scenarios of the hydrological system and the areas with potential risks of future dangerous groundwater table depths in a Brazilian Cerrado area. Similarly, neural network methods have been used by many authors. Back Propagation Neural Network (BPNN) and Radial Basis Function Network (RBFN) models (Ghose et al. 2010) were considered for this study in order to overcome the difficulties of identifying the nonlinear model structure and estimating the associated parameters. Both models use historical groundwater level records and related hydro-meteorological data to simulate groundwater table fluctuations in the study area.

The Hybrid Artificial Neural Network-Geostatistics Model was studied by Nourani et al. (2011) to comprehend the spatiotemporal variations of the groundwater level for the management of groundwater in coastal areas. In the last few years, the Artificial Neural Network (ANN) has been used to forecast groundwater levels in several studies (e.g., Chitsazan et al. 2015; Daliakopoulos et al. 2005; Djurovic et al. 2015; Mohanty et al. 2015; Sun et al. 2016; and Tsanis et al. 2008). In addition, Varouchakis (2016) used autoregressive exogenous variable (ARX) model parameters by means of the Kalman filter adaptation algorithm (KFAA) for a 3-year forecast on the island of Crete. Wunsch, et al. (2018) applied nonlinear autoregressive networks with exogenous input (NARX) to obtain groundwater level forecasts for several wells in southwest Germany. Amaranto et al. (2018) hypothesize that data-driven modeling (e.g., ANNs, support vector machines, random forests, genetic programming and extreme machine learning) can improve naïve and autoregressive simulations of groundwater tables.

To predict future events using traditional tools without losing accuracy and precision in the forecast, this study investigates two well-known method—one geostatistical simulation (SGS) and one time series forecaster (an ARIMA model), to address the problem of predicting groundwater table depth in the Bauru Aquifer System (BAS). The flowchart for the application of both methods and the comparisons of results is shown in Figure 1.

Time Series Modeling

Time series are analyzed to understand the past and predict the future, enabling managers or policymakers to make properly informed decisions. A time series analysis quantifies the main features of a set of data and its random variation. This feature of time series analysis, combined with improved computing power, has made time series methods widely applicable in government, industry, and commerce (Cowpertwait and Metcalfe 2009). Shumway and Stoffer (2011) explained that the primary objective of time series analysis is to develop mathematical models that provide plausible descriptions of sample data. In order to provide a statistical setting for describing the features of data that seemingly fluctuate in a random fashion over time, it is assumed that a time series can be defined as a collection of
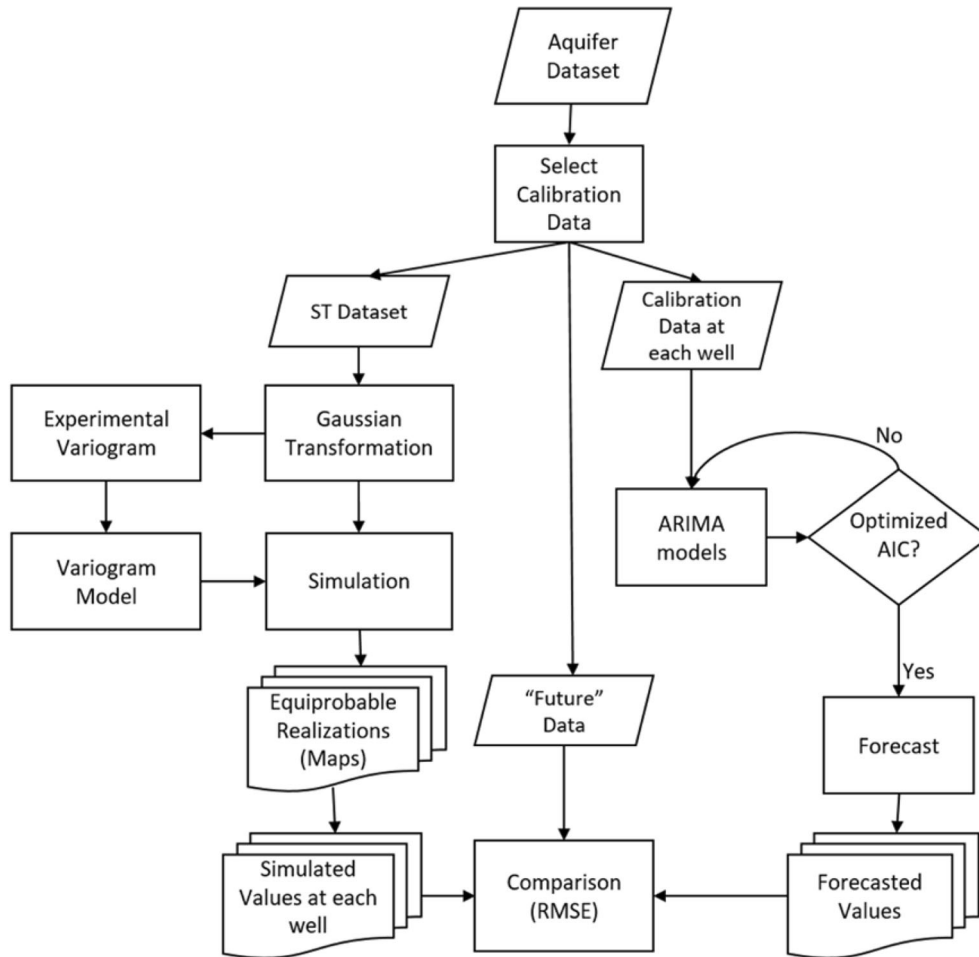
Figure 1. Flowchart of the methodology of the study and comparison of results.

random variables indexed according to the order in which they are obtained in time.

The basic steps in time series modeling and forecasting, as described by Montgomery et al. (2015), are to plot the time series and determine its basic features, such as whether trends or seasonal behavior or both are present. Then, look for possible outliers or any indication that the time series has changed with respect to its basic features (such as trends or seasonality) over the period been studied. Next, eliminate any trend or seasonal components, either by differencing or by fitting an appropriate model to the data. Moreover, consider using data transformations, particularly if the variability in the time series seems to be proportional to the average level of the series. Montgomery et al. (2015) further explained that the objective of these operations is to produce a set of stationary residuals. Finally, de-

velop a forecasting model for the residuals. It is not unusual to find several plausible models, and so additional analysis will have to be performed to determine the best one for deployment. Sometimes, potential models can be eliminated based on their degree of fit to the historical data. It is unlikely that a model that fits poorly will produce good forecasts.

Montgomery et al. (2015) demonstrate that one should validate the performance of a model (or models) from the previous step. This will probably involve some type of split-sample or cross-validation procedure. The objective of this step is to select a model to use in forecasting. Furthermore, it is interesting to compare the differences between the original time series and the values that would be forecast by the model on the original scale. Details about traditional time series methods can be found in Chatfield (2003), De Gruijter et al. (2006), Prado

and West (2010), Brockwell and Davis (2010), Cressie and Wikle (2011) and Box et al. (2016).

Box and Jenkins (1970), Newbold et al. (1993), and Janacek (2001) define the ARIMA equation as:

$$\Phi(B)(1-B)^d.X_t = \theta(B).\varepsilon_t, \tag{1}$$

where Φ and θ are polynomials in the back-shift operator B and ε_t is white noise. Let p denote the number of parameters in the autoregressive operator Φ, d the degree of differencing and q the number of parameters in the moving average operator θ; this is designated as the ARIMA (p, d, q) model.

In order to evaluate and choose the best model, there are a few fundamental procedures to identify the orders of ARIMA models. The first one is the Akaike information criterion (AIC), which is the most common and was described by Akaike (1974) as

$$\text{AIC} = -2\log(L) + 2k, \tag{2}$$

where L is the maximum likelihood and k is the number of independently adjusted parameters within the model.

Woodward et al. (2011) explained that a problem with using AIC is that as the realization length increases, AIC tends to overestimate the model orders. Several authors have considered modifications to AIC to adjust for this problem. The most common alternatives to AIC are the Bayesian information criterion (BIC) and the corrected Akaike information criterion, proposed by Schwarz (1978) and Hurvich and Tsai (1989), respectively.

Geostatistical Simulation

According to Chilès and Delfiner (1999), conditional simulations are qualitatively useful for obtaining realistic pictures of spatial variability. Quantitatively, they are the tool of choice for evaluating the impact of spatial uncertainty on the results of complex procedures, such as numerical modeling of a dynamic system or economic optimization of the development of a natural resource. The goal pursued here is not to reproduce the genetic mechanisms that generated the observed phenomenon but simply to mimic its spatial variations as realistically as possible. The results of a simulation must always be carefully considered and checked against the background of the application. More details about geostatistical simulation and its applications, mathematical proofs, and parameters can be found in Journel and Huijbregts (1978), Lantuéjoul (2002), Goovaerts (1997), and Deutsch and Journel (1998), among others. Furthermore, Nunes and Almeida (2010) presented the parallelized algorithm of sequential Gaussian, indicator, and direct simulations.

Olea (1999) pointed out that geostatistics circumvents the problem of smoothing in kriging with stochastic simulation. Unfortunately, simulation is only a trade-off, not a perfect solution; stochastic realizations are not error-free renditions of reality. Lacking a perfect solution, the choice between kriging and simulation must be decided based upon which is more relevant for each specific application: minimum local estimation errors in a mean-square sense or correct spatial continuity. Deutsch and Journel (1998) explained that in simulations, reproduction of global features (texture) and statistics (histogram, covariance) take precedence over local accuracy. Kriging provides a set of local representations where local accuracy prevails. Simulation provides an alternative global representation, where reproduction of patterns of spatial continuity prevails. Finally, Goovaerts (1997) clarified that the set of alternative realizations generated by stochastic simulation provides a measure of uncertainty about the spatial distribution of attribute values.

One method used in this study is SGS performed with SGeMS (Stanford Geostatistical Modeling Software), whose algorithm was described by Remy et al. (2011) as follows: transform the data distribution into a Gaussian distribution with 0 mean and variance equal to 1 ($N[0,1]$), and then define a random path to visit each node of the grid. At each node, get the conditioning data consisting of neighboring original hard data and previously simulated values and estimate the local conditional cdf as a Gaussian distribution with mean given by simple kriging (Eqs. 3 and 4) and variance by simple kriging variance (Eq. 5). Finally, draw a value from that Gaussian ccdf (the theory guarantees variogram reproduction only when simple kriging is used) and add the simulated value to the dataset. Repeat for all realizations and back transform the Gaussian simulated field into the data space. The algorithm calls for the variogram of the normal score, not of the original data. Only the normal score variogram is

guaranteed to be reproduced within ergodic fluctuations.

Journel (1989) gave the simple kriging equation as:

$$Z_0^* = m_0 + \sum_{i=0}^{n} \lambda(Z_i - m_i), \qquad (3)$$

where m_0 is the expected value, m_i is the mean of the samples, Z_i is the dataset, and λ_i is the kriging weight calculated by the linear equation system as:

$$K\lambda = k, \qquad (4)$$

where K is the spatial covariance matrix of the sampled values and k is the spatial covariance matrix of the data-to-unknown. The kriging variance is:

$$\sigma_{SK}^2 = C_{00} - \sum_{i=1}^{n} \lambda_i k. \qquad (5)$$

CASE STUDY: FORECASTING GROUNDWATER TABLE DEPTHS IN A BAURU AQUIFER SYSTEM AREA

From 2013 to 2016, São Paulo State (SP), Brazil, passed through two marked periods of climatic anomalies, facing one of the worst droughts ever recorded (Coelho et al. 2016) and later the effects of El Niño South Oscillation (ENSO) phenomena, both of which had a direct impact on water resources. A monitoring network of groundwater table depths and an automatic climatological station installed at the Ecological Station of Santa Barbara (EEcSB) in the municipality of Águas de Santa Barbara (SP), Brazil, captured the response of groundwater levels to these climatic inputs. The area is covered by sediments making up the BAS, which is highly vulnerable due to its superficial levels. In addition to the dangers of contamination caused by anthropogenic activities, the BAS is sensitive to seasonal climatic variations due to its dependence on direct recharge by precipitation. Such information allows us to verify the temporal and spatial dynamics of characteristics (Von Asmuth et al. 2008) such as the response time of an aquifer to a precipitation event; that is, how long a groundwater intake in the system will influence the groundwater table oscillation (Hocking and Kelly 2016; Manzione et al. 2017).

Study Area: Ecological Station of Santa Barbara (EEcSB)

Melo and Durigan (2011) reported that the EEcSB was established in 1984 and is located in the municipality of Águas de Santa Bárbara, covering 2712 ha of an ecologically important land reserve located in the Middle Paranapanema watershed. The main vegetation type in the area is Cerrado, and this area lies on the border of another important biome, the Atlantic Forest. The zone created by this intersection is of great importance because such transition zones are areas with high biodiversity. Moreover, Melo and Durigan (2011) also reported that the temperature averages range from 16.9 to 24.3 °C from the coldest to the hottest months, respectively, the pluviosity averages range from 44 mm in August to 206 mm in December, and the topography includes hills with elevations between 600 and 680 m. The geology is basically composed of sandstones of the Adamantina and Marilia Formations of the Bauru Group and extrusive igneous rocks and basalts of the Serra Geral Formation of the São Bento Group.

The Bauru aquifer is important for the water supply in São Paulo State because of its expansive territorial distribution (106,996 km^2), medium flow rates and easy access (DAEE 2013). Moreover, it has a mean saturated thickness of 75 m and reaches 300 m in the Residual Plateau of Marília, conditioned by the surface morphology and the rocky substrate represented by the basalts of the Serra Geral Formation. It is recharged by rainfall, and the exploitable flow varies from 10 to 120 m^3/h, with high potential zones located near the Paraná River (CETESB 2013). For these reasons, the Bauru Aquifer is used extensively in public water supply systems to provide drinking water (Santos and Bonotto 2014).

The data were acquired at 49 water wells near springs (Fig. 2) at the EEcSB—BAS, in the countryside of São Paulo state, Brazil. The groundwater level has been measured since September 5th, 2014, and this study investigated a hydrological year (until September 30, 2015). To properly compare the results, the predictions were done on the next four sampled dates: October 16, October 29, November 13, and December 3, 2015.
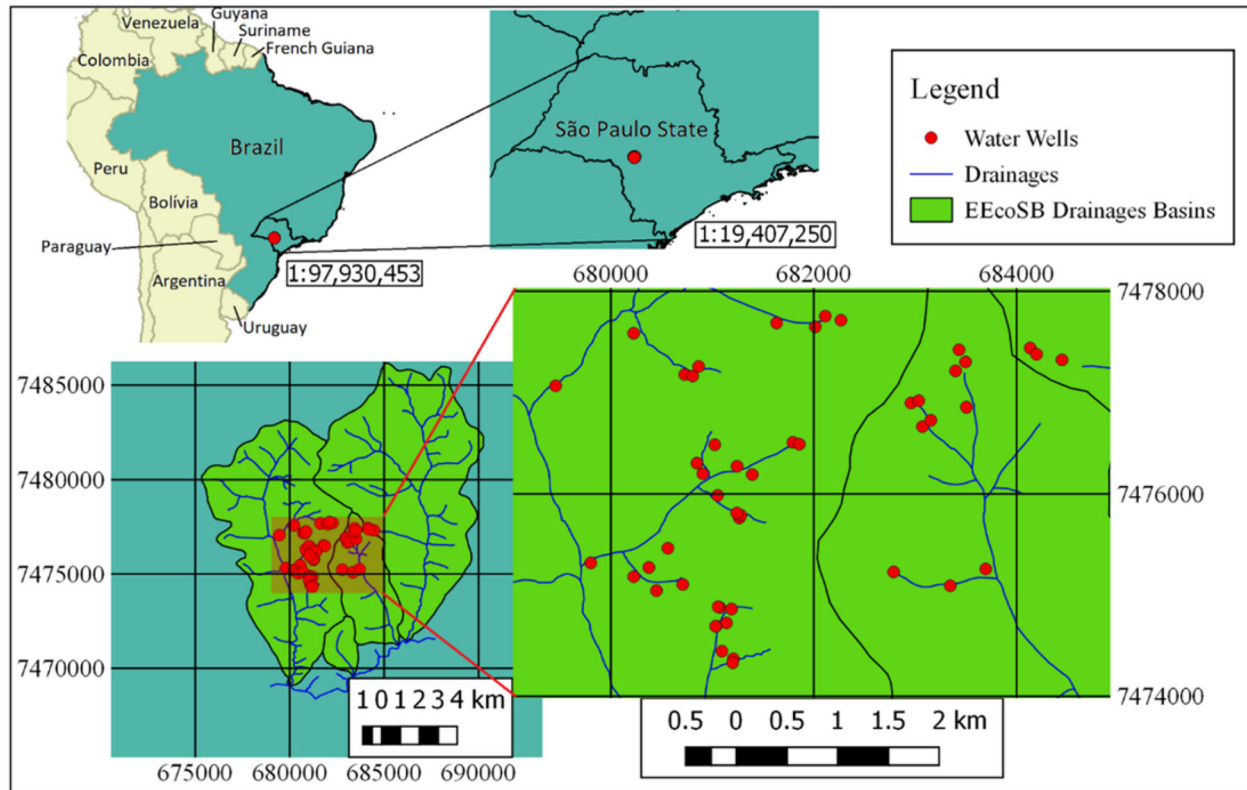
Figure 2. Location of the water wells inside the Bauru Aquifer System (BAS) in the countryside of São Paulo state, Brazil.

Time Series Modeling (ARIMA)

To accomplish this study, the R package (R Core Team 2016) called 'forecast', presented by Hyndman (2017) and Hyndman and Khandakar (2008), was used to perform the time series analysis and forecast the future data as an automatic AR-IMA model (Eq. 1). One function was set for each well, and its parameters were automatically adjusted using the minimization of the AIC (as shown in Eq. 2) to fit the original data variability curve with the trend, seasonal, cyclical and irregular components. Moreover, the same parameters were used to forecast the next steps in time. Figure 3 shows the six randomly selected examples of the modeled time series used. Moreover, a total of 16 water wells were measured beginning in the middle of 2015; their forecasted line closely reproduces the pattern of their data (Fig. 3c). Although they have few measured values, they seem to be well reproduced by the time series.

To evaluate the model calibration, the error metric root-mean-square error (RMSE) of the water wells is presented in the form of a histogram (Fig. 4). The RMSE is a measure for goodness of model fit; its values range from 0.002 to 0.158 m with an average of 0.052 m. Figure 4 also shows the histogram of each optimized AIC, with an average of − 131.194, a maximum value of − 3.412 and a minimum value of − 307.582. The AIC is a measure of the likelihood and the number of parameters. When comparing models for the same time series, it values simplicity (a lower number of parameters is better), as shown in Eq. 2.

Geostatistical Simulation (SGS)

SGS is considered to be a framework for all spatial points on all dates. The spatial distribution map uses the north and east as the Y- and X-axes, respectively, and, aiming for computational simplicity, time is considered to be a dimension on the Z-axis. So, it is considered to be a simple spatiotemporal cube for the geostatistical simulation. In terms of geostatistical covariance models, this con-
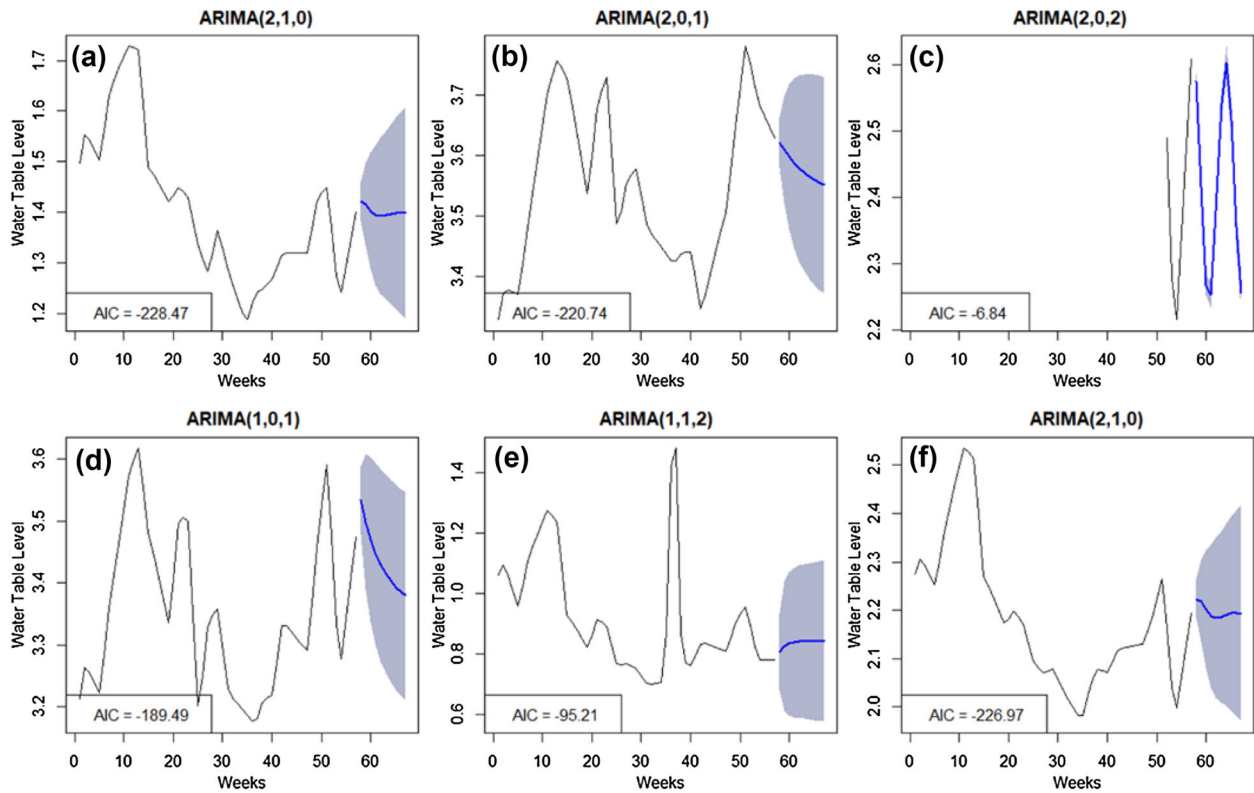
Figure 3. Six selected examples of ARIMA models. The black solid line is the original data, the forecasted blue line starts after 55 weeks, and the gray area is the confidence level of 80%.
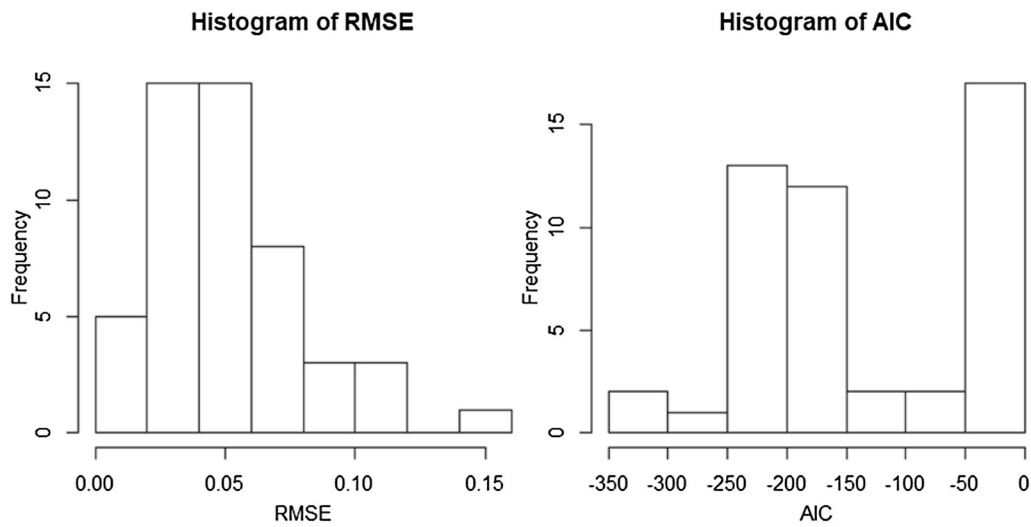

Figure 4. Histograms of error measures obtained from model calibrations: RMSE (left) and AIC (right).

sideration is equivalent to the spatiotemporal metric covariance model (Dimitrakopoulos and Luo 1994) when the covariance is computed through the Euclidian distance in space (h) and time (u) with the correction factor $c = 1$. The metric model is

$$C(h, u) = C\left(\sqrt{\|h\|^2 + c|u|^2}\right). \quad (6)$$

To proceed with the SGS, the data were transformed into a Gaussian distribution ($N[0,1]$) by means of the normal score, which guarantees that any distribution will be converted to Gaussian. Then, the experimental variogram was computed and the model was fitted with two spherical structures with a nugget effect (C_0) of 0.013. The spherical variogram model was chosen because it presented the best fit among the most common validated variogram models (spherical, cubic, exponential and Gaussian). The first structure (C_1) has a spatiotemporal variance of 0.118, a maximum and medium range of 750 m in space and a minimum range of 185 days in time. The second structure (C_2) has a spatiotemporal variance of 0.869 and maximum and medium ranges of 750 m in space, and the time dimension axis has already reached its sill. The anisotropy ellipsoid has its major and medium axes at a spatial domain and its minimum at the temporal domain. Figure 5 shows that the given model (the black line with parameters in red) fits the experimental variogram (red) well. Moreover, the complete model equation is

$$\gamma(h,u) = 0.013 + 0.118\left(\frac{3}{2}\left(\frac{(h,u)}{\sqrt{185^2 + 750^2}}\right)\right.$$
$$\left. -\frac{1}{2}\left(\frac{(h,u)}{\sqrt{185^2 + 750^2}}\right)^3\right)$$
$$+ 0.869\left(\frac{3}{2}\left(\frac{(h,u)}{750}\right) - \frac{1}{2}\left(\frac{(h,u)}{750}\right)^3\right)$$

This simulation generates as many realizations as desired, and all of them have the same probability of agreeing with reality. In our study, 100 realizations were computed—the number of simulations needed depends on the parameter distribution, and typically 100 simulations are enough to assess the range of possibilities (Chilès and Delfiner 1999). The simulation was computed using up to three sampled points and two previously simulated points. Such a low number of points was used in the simulation because of the high probability of considering only the points at the same location and different dates

and because it is desirable to use only the closest points, both in time and space, without biasing the data. In addition, the neighborhood search ratio was defined as the same distance in each variogram range. The simulation was performed in the geographic domain at the four future dates. The SGeMS was used for the geostatistical method; it includes data transformation, variogram modeling, simulation and statistics (mean, P(10), and P(90)). Therefore, the simulated maps were plotted (Fig. 6) using R with the 'raster' package (Hijmans 2017). The figure shows the simulated groundwater table level maps on October 16, October 29, November 13, and December 3 and the ARIMA forecasted values at the wells locations. It seems that the results are similar at the map scale, but their errors presented in the next section are not.

EVALUATION OF MODEL PERFORMANCE

The simulated values at the water well locations and at the same dates for which there are observed data were selected for evaluation of model performance. The data selected to predict the four "future" sampled points have a spatial distribution with north and east plotted on the Y- and X-axes, respectively—as shown on the map in Figure 1—and a temporal factor that was considered on the Z-axis. To compare the methods and the sampled data visually, a line chart was created with all the water wells on the X-axis and the groundwater level on the Y-axis. The sequence presented on the water well axis is arbitrary and does not represent the connectivity and continuity of the groundwater table. Thus, it is possible to see the evolution of both prediction methods through time and to compare these methods on the same date.

Figure 7 shows the predictions for October 16 obtained using the ARIMA and SGS methods. We see that both approaches did well at this step; however, the time series presented a smaller uncertainty range. For the second simulated date (October 29th), as shown in Figure 8, both methods have good expected values to represent the sampled data; however, both uncertainty ranges are larger compared to those on the first date. The forecasting and simulation for the third date (November 13th) are presented in Figure 9. For most of the wells, both methods kept the sampled data within their certainty ranges, although the sampled data are closer to
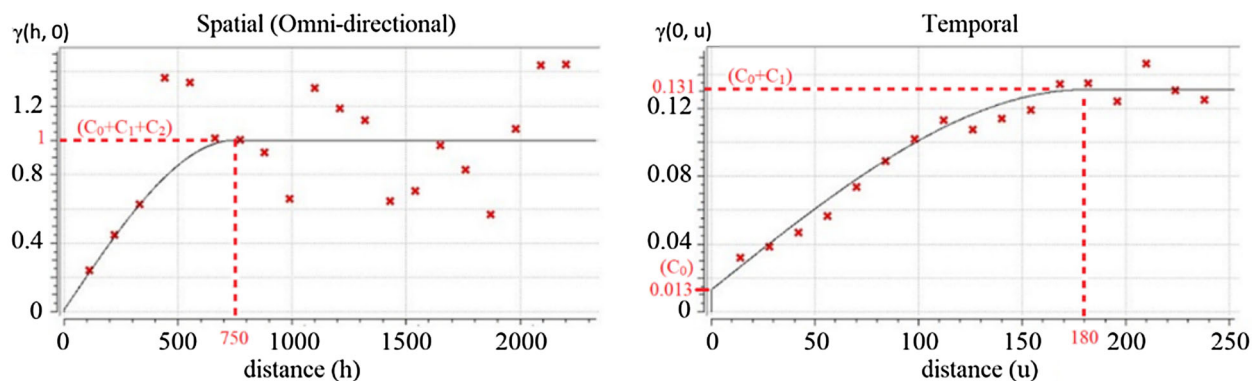
Figure 5. Variogram of the original data, considering X (east) and Y (north) as spatial factors and Z (time) as a temporal factor.
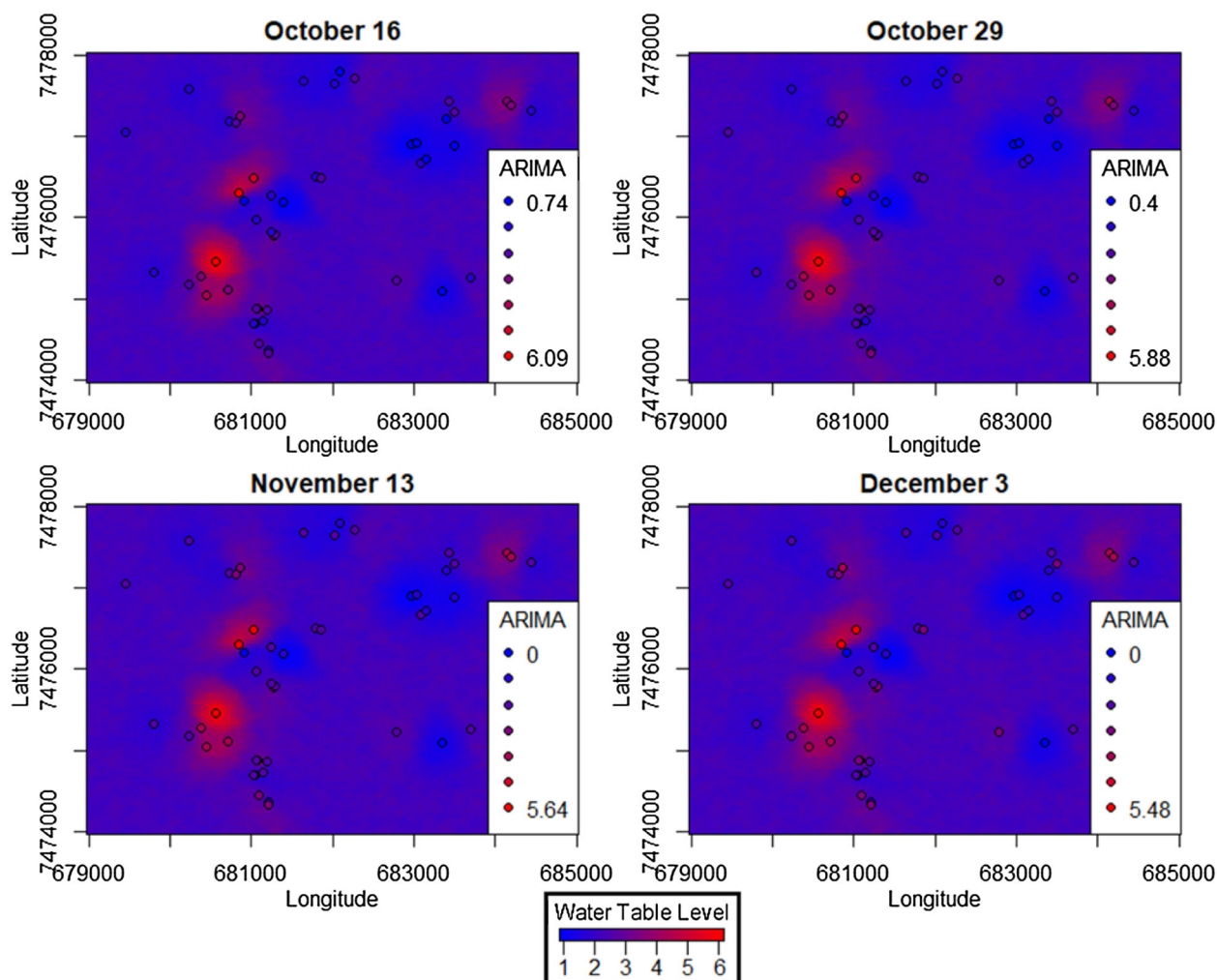

Figure 6. Simulated groundwater table level at all domains; the circles are the values of the ARIMA model forecasts.
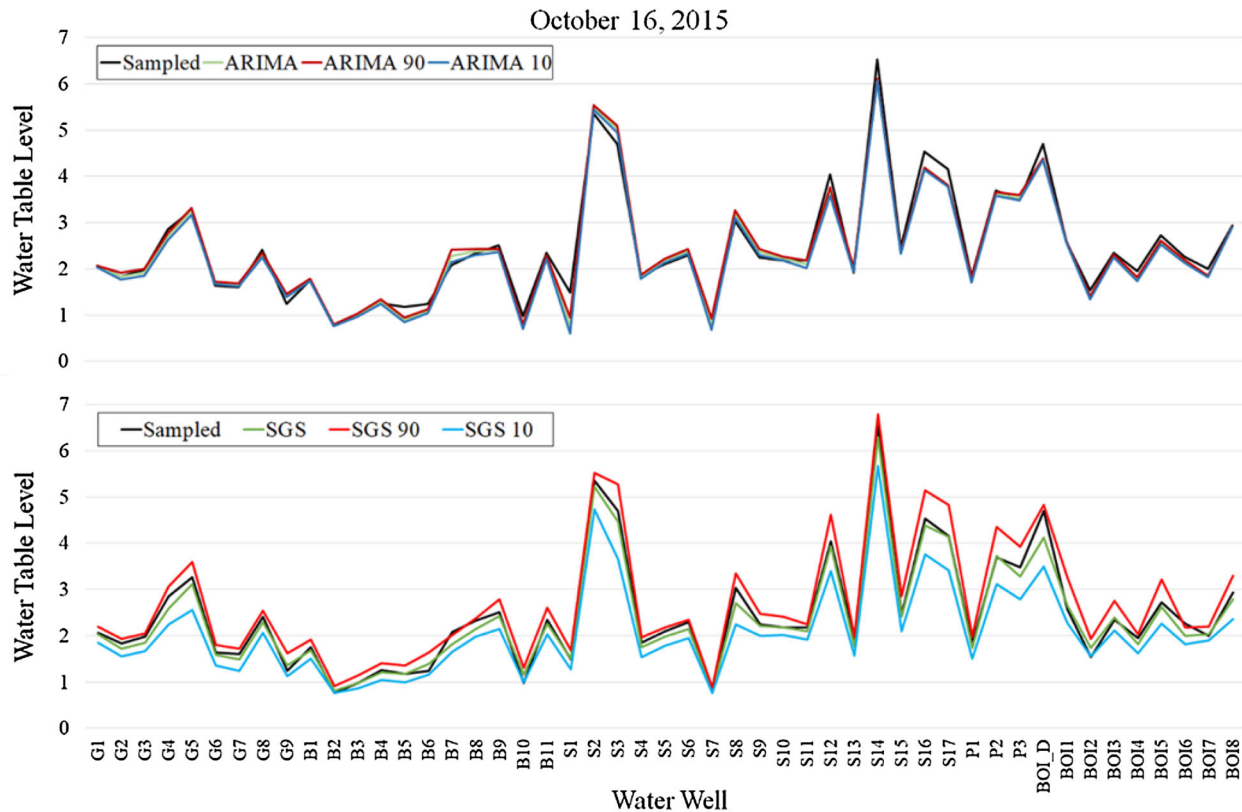
Figure 7. Results obtained with the ARIMA and SGS methods for October 16, 2015. The solid black line represents the sampled points at ''future'' dates, the green line represents the estimated values, and the blue line and the red line represent P(10) and P(90), respectively.

P(10) obtained from the forecasting and simulation approaches. Figure 10 illustrates the forecasting and simulation methods on the fourth prediction date, December 3, where the prediction results of both methods were less trustworthy compared to those on the first three dates. However, the simulation performed better since its P(10) is closer to the sampled data than that of ARIMA. To evaluate the results quantitatively, the average error was calculated by subtracting the calculated value from the observed value. Figure 11 shows the errors per estimated date, and it corroborates the visual analysis of Figures 7, 8, 9 and 10.

Considering the problem of predicting the groundwater table level, Knotters and Bierkens (2001) showed RMSE = 0.34 m for the regionalized autoregressive exogenous variable (RARX) model and RMSE = 0.22 m with DEM as auxiliary information. Bierkens et al. (2001) showed that for the RARX with Kalman filter present, the RMSE = 0.30 m, and with auxiliary information (DEM), the

RMSE = 0.14 m. Bierkens (2001) considered a stochastic differential equation to predict and kriging with a trend to interpolate with RMSE = 0.27 m. Moreover, Omran (2016) showed an average RMSE of 0.16 m for time series modeling. Manzione et al. (2017) applied the PIRFICT model and obtained RMSE ranging from 0.05 to 1.16 m. Wunsch et al. (2018) demonstrated an average RMSE = 0.604 m for NARX model validation, and average RMSEs for 1-, 4-, 12-, and 26-week forecasts of 0.04, 0.17, 0.20, and 0.29 m, respectively. Currently, ANN is widely used for groundwater level research and this method provides low error rates with average RMSE = 0.15 m (Djurovic et al. 2015), average RMSE = 0.09 m for a 7-day forecast (Sun et al. 2016), and average RMSEs ranging from 0.24 m (1 week) to 0.48 m (4 weeks) in the training and from 0.41 m (1 week) to 0.61 m (4 weeks) in the forecasting (Mohanty et al. 2015). Moreover, Amaranto et al. (2018) made a 5-month analysis, considering multiple data-driven models, and com-
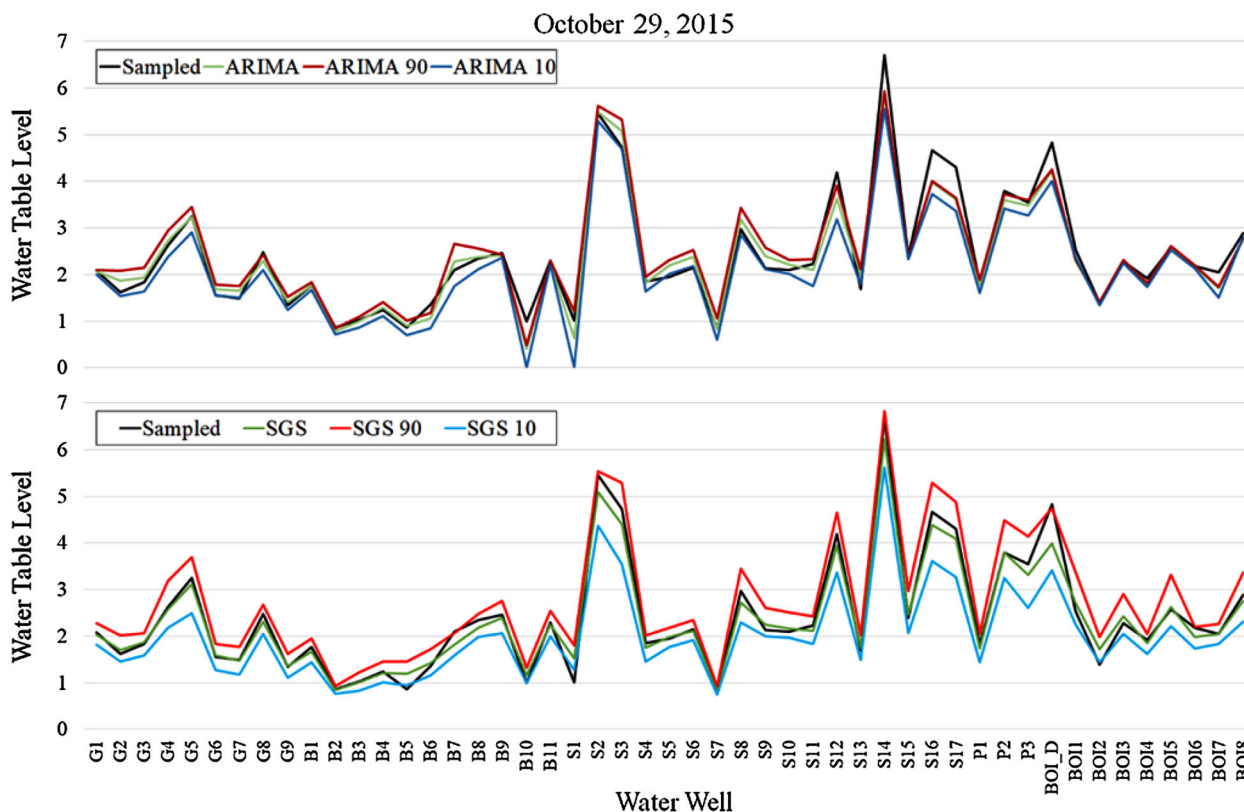
Figure 8. Results obtained with the ARIMA and SGS methods for October 29, 2015. The solid black line represents the sampled points at "future" dates, the green line represents the estimated values, and the blue line and the red line represent P(10) and P(90), respectively.

pared the rising limb with RMSE from 0.1 to 0.32 m to the falling limb with RMSE from 0.18 to 0.45 m.

Table 1 shows that considering a spatiotemporal SGS is not a bad idea, due to the fact that its RMSE is among the best ones for a short-term forecast. As expected, the RMSE worsens when the predictions are significantly different from the calibration data. The SGS is more accurate for short-term forecast, but errors for the 2-month forecast are equivalent to those obtained by ARIMA. The analysis of the highest (P(10)) and lowest (P(90)) level errors shows that ARIMA models are always more precise and have lower average RMSE. However, after some steps, the inferior scenario of the simulation (SGS P(10)) gets closer to the reality due to the arid period, resulting in lower RMSE. Furthermore, one reason why the forecast is very poor after a few time steps may be the anomalous period: it was one of the driest periods ever recorded in São Paulo State, and it was in the middle of an El Niño

event. However, the short-term prediction is demonstrated to be very precise and accurate for both methods.

Predicting the potentiometric cartography is important for knowing the flow movement as well as the spatial variance of hydrogeology properties. Moreover, it can be used directly for public administration (soil use and management) because any anthropogenic action reflects the potential energy, and it is possible to map where it is a good idea to invest on water facilities and where it is not a bad idea to have gas stations, industries and all potentially polluting sources, e.g., to analyze soil use and manage it properly, it is necessary to consider the whole region. Moreover, although ecological stations are environmentally protected, any change in groundwater level can impact their biomes because some types of vegetation need higher groundwater table levels than others. Thus, not only is pollution bad, but its effect on the groundwater table can be dangerous to environmental equilibrium. Near the
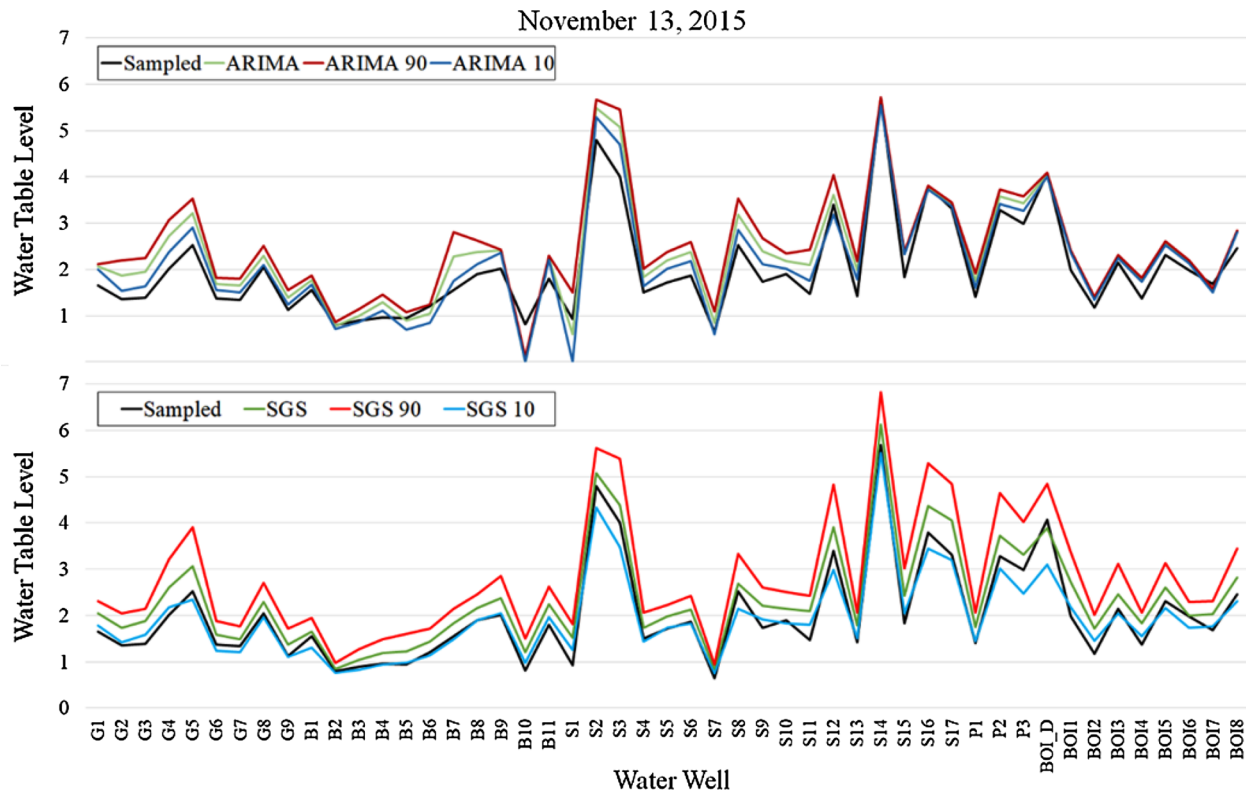
Figure 9. Results obtained with the ARIMA and SGS methods for November 13, 2015. The solid black line represents the sampled points at "future" dates, the green line represents the estimated values, and the blue line and the red line represent P(10) and P(90), respectively.

study area, there are eucalypt farms, and it is possible that they are influencing the groundwater levels inside the ecological station. Furthermore, our predictions were done in an anomaly year: It was one of the driest periods ever recorded in São Paulo State, probably because of the El Niño influence, so seasonality is not fully represented. Therefore, these models should consider the lower-frequency cyclicity of El Niño. Despite all of these obstacles, both models performed reasonably well for a short-term forecast.

The case study is important to define the best approach to predict the groundwater depth inside the EEcSB in the countryside of São Paulo state, Brazil. Despite the short time series and the fact that the period was in one of the driest years recorded in the area and was during El Niño, the comparison between the results of SGS and ARIMA shows that both methods worked well when forecasting just a few days in the future. At the second and third prediction dates, the expected values of the simulation and the time series were

similar, with a wider range of possibilities in the simulation approach. In the long-term (more than 2 months), the groundwater table level is decreasing because of the water scarcity in this period. As a result, the sampled values tend to the lower extreme of the predicted/forecasted scenario (P(10)) and only the simulation contemplate this anomalous event.

The time series is an extrapolation method, so it is expected to work better than an interpolation method, such as SGS. The advantage of SGS in the long term was that its possible range was wide enough to cover the actual data. For distances beyond the neighborhood search, the simulation will run only with previously simulated data and any pattern will be extrapolated. The cost of the long-term extrapolation is that SGS tends to honor global mean and variance. The major limitation of this kind of prediction is that the back transformation—i.e., returning from the Gaussian distribution to the original data—will force the final distribution to be the same as the sampled data, including the mini-
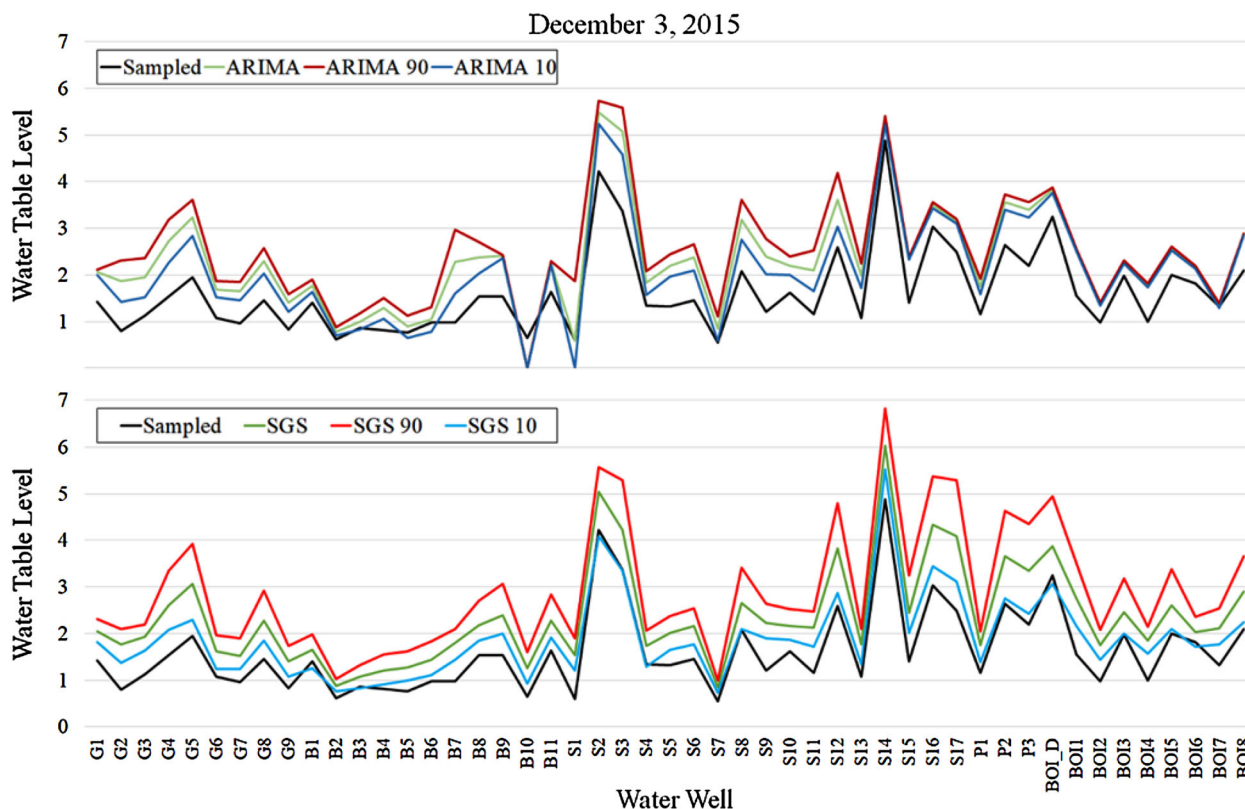
Figure 10. Results obtained with the ARIMA and SGS methods for December 3, 2015. The solid black line represents the sampled points at "future" dates, the green line represents the estimated values, and the blue line and the red line represent P(10) and P(90), respectively.
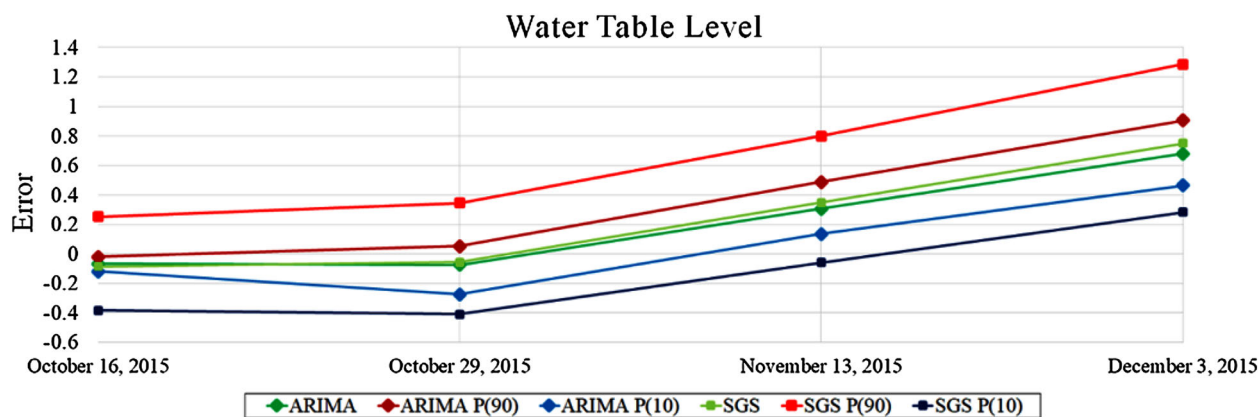

Figure 11. Groundwater level errors from the methods applied to groundwater series from EEcSB.

mum, mean and maximum values. In addition, the comparison of analysis for longer periods of time may be necessary to better predict long-term groundwater evolution and improve time series predictions.

CONCLUSION

The analysis described showed that both SGS and ARIMA predicted reasonably well for a short-term forecast. The ARIMA model performed

Table 1. Root-mean-square error (RMSE) of the forecasted dates

	October 16, 2015	October 29, 2015	November 13, 2015	December 3, 2015
ARIMA	0.20 m	0.28 m	0.44 m	0.80 m
ARIMA P(10)	0.22 m	0.42 m	0.32 m	0.60 m
ARIMA P(90)	0.18 m	0.31 m	0.62 m	1.05 m
SGS	0.16 m	0.22 m	0.39 m	0.81 m
SGS P(10)	0.31 m	0.54 m	0.25 m	0.37 m
SGS P(90)	0.47 m	0.41 m	0.87 m	1.39 m

favorably with regard to groundwater table depth forecasts for short-term periods because of their good accuracy and higher precision. In contrast, the SGS is more accurate and its wide range of possible scenarios captured anomalous events such as the severe drought monitored during this study. The choice between these models is a precision and accuracy trade-off. In order to implement a system that updates the dataset and forecast automatically, the ARIMA models are a great option because they can be modeled automatically using AIC; moreover, their results reproduce the trend and seasonality patterns of the data with satisfactory precision and accuracy. However, considering all wells as a spatiotemporal framework, SGS can generate the groundwater table level map. However, this method does not support automatic modeling for the variogram, which precludes self-monitoring system.

REFERENCES

Akaike, H. (1974). A new look at the statistical model identification. *IEEE Transactions on Automatic Control, 19*(6), 716–723. https://doi.org/10.1109/TAC.1974.1100705.

Amaranto, A., Munoz-Arriola, F., Corzo, G., Solomatine, D. P., & Meyer, G. (2018). Semi-seasonal groundwater forecast using multiple data-driven models in an irrigated cropland. *Journal of Hydroinformatics*. https://doi.org/10.2166/hydro.2018.002. **(in press)**.

Bierkens, M. F. P. (2001). Spatio-temporal modelling of the soil water balance using a stochastic model and soil profile descriptions. *Geoderma, 103*(1–2), 27–50. https://doi.org/10.1016/S0016-7061(01)00068-4.

Bierkens, M. F. P., Knotters, M., & Hoogland, T. (2001). Space-time modeling of water table depth using a regionalized time series model and the Kalman Filter. *Water Resources Research, 37*(5), 1277–1290. https://doi.org/10.1029/2000WR900353.

Box, G. E. P., & Jenkins, G. M. (1970). *Time series analysis: Forecasting and control*. San Francisco: Holden-day.

Box, G. E. P., Jenkins, G. M., Reinsel, G. C., & Ljung, G. M. (2016). *Time series analysis: Forecasting and control* (5th ed.). Hoboken: Wiley.

Brockwell, P. J., & Davis, R. A. (2010). *Introduction to time series and forecasting* (2nd ed.). New York: Springer.

Castrignanò, A., & Buttafuoco, G. (2004). Geostatistical stochastic simulation of soil water content in a forested area of south Italy. *Biosystems Engineering, 87*(2), 257–266. https://doi.org/10.1016/j.biosystemseng.2003.11.002.

CETESB. (2013). *Qualidade das águas subterrâneas do estado de São Paulo 2010–2012*, Companhia Ambiental do Estado de São Paulo. http://cetesb.sp.gov.br/wp-content/uploads/sites/42/2013/11/aguas_sub_2012.pdf. Accessed July 13, 2017.

Chatfield, C. (2003). *The analysis of time series: An introduction*. Boca Raton: CRC Press Company.

Chilès, J. P., & Delfiner, P. (1999). *Geostatistics: Modeling spatial uncertainty*. Hoboken: Wiley.

Chitsazan, M., Rahmani, G., & Neyamadpour, A. J. (2015). Forecasting groundwater level by artificial neural networks as an alternative approach to groundwater modeling. *Journal of the Geological Society of India, 85*(1), 98–106. https://doi.org/10.1007/s12594-015-0197-4.

Coelho, C. A. S., Cardoso, D. H. F., & Firpo, M. A. F. (2016). Precipitation diagnostics of an exceptionally dry event in São Paulo, Brazil. *Theoretical and Applied Climatology, 125*(3–4), 769–784. https://doi.org/10.1007/s00704-015-1540-9.

Cowpertwait, P. S. P., & Metcalfe, A. V. (2009). *Introductory time series with R*. New York: Springer.

Cressie, N., & Holan, S. H. (2011). Editorial: Special issue on time series in the environmental sciences. *Journal of Time Series Analysis, 32*(4), 337–338. https://doi.org/10.1111/j.1467-9892.2011.00739.x.

Cressie, N., & Wikle, C. K. (2011). *Statistics for spatio-temporal data*. Hoboken: Wiley.

DAEE. (2013). *Águas subterrâneas no Estado de São Paulo*. Diretrizes de Utilização e Proteção, Departamento de Águas e Energia Elétrica do Estado de São Paulo. http://www.daee.sp.gov.br/acervoepesquisa/Atlas%20-%20%C3%81guas%20Subterr%C3%A2neas%20(DAEE-LEBAC).pdf. Accessed July 13, 2017.

Dagan, G. (2002). An overview of stochastic modeling of groundwater flow and transport: From theory to application.

EOS (Transactions American Geophysical Union), 83(53), 621–625. https://doi.org/10.1029/2002EO000421.

Daliakopoulos, I. N., Coulibaly, P., & Tsanis, I. K. (2005). Groundwater level forecasting using artificial neural networks. *Journal of Hydrology, 309,* 229–240. https://doi.org/10.1016/j.jhydrol.2004.12.001.

De Gruijter, J. J., Brus, D. J., Bierkens, M. F. P., & Knotters, M. (2006). *Sampling for natural resource monitoring*. Berlin: Springer.

Deutsch, C. V., & Journel, A. G. (1998). *GSLIB: Geostatistical software library and user's guide* (2nd ed.). New York: Oxford University Press.

Dimitrakopoulos, R., & Luo, X. (1994). Spatiotemporal modeling: Covariances and ordinary kriging system. In R. Dimitrakopoulos (Ed.), *Geostatistics for the next century* (pp. 88–93). Dordrecht: Springer.

Djurovic, N., Domazet, M., Stricevic, R., Pocuca, V., Spalevic, V., Pivic, R., et al. (2015). Comparison of groundwater level models based on artificial neural networks and ANFIS. *The Scientific World Journal.* https://doi.org/10.1155/2015/742138.

Ghose, D. K., Panda, S. S., & Swain, P. C. (2010). Prediction of water table depth in western region, Orissa using BPNN and RBFN neural networks. *Journal of Hydrology, 394*(3–4), 296–304. https://doi.org/10.1016/j.jhydrol.2010.09.003.

Goovaerts, P. (1997). *Geostatistics for national resources evaluation*. New York: Oxford University Press.

Hatch, C. E., Fisher, A. T., Revenaugh, J. S., Constantz, J., & Ruehl, C. (2006). Quantifying surface water–groundwater interactions using time series analysis of streambed thermal records: Method development. *Water Resources Research, 42,* W10410. https://doi.org/10.1029/2005WR004787.

Hijmans, R. J. (2017). 'raster': Geographic data analysis and modeling. R Package (version 2.6-7). https://CRAN.R-project.org/package=raster Accessed May 13, 2018.

Hocking, M., & Kelly, B. F. J. (2016). Groundwater recharge and time lag measurement through Vertosols using impulse response functions. *Journal of Hydrology, 535,* 22–35. https://doi.org/10.1016/j.jhydrol.2016.01.042.

Hurvich, C. M., & Tsai, C. L. (1989). Regression and time series model selection in small samples. *Biometrika, 76*(2), 297–307. https://doi.org/10.1093/biomet/76.2.297.

Hyndman, R. J. (2017). 'forecast': Forecasting functions for time series and linear models. R package version 8.0, http://github.com/robjhyndman/forecast. Accessed July 12, 2017.

Hyndman, R. J., & Khandakar, Y. (2008). Automatic time series forecasting: The forecast package for R. *Journal of Statistical Software, 27*(3), 1–22. https://doi.org/10.18637/jss.v027.i03.

Janacek, G. (2001). *Practical time series*. London: Hodder Education Publishers.

Journel, A. G. (1989). *Fundamentals of geostatistics in five lessons*. Washington: American Geophysical Union.

Journel, A. G., & Huijbregts, Ch J. (1978). *Mining geostatistics*. San Diego: Academic Press.

Knotters, M., & Bierkens, M. F. P. (2001). Predicting water table depths in space and time using a regionalized time series model. *Geoderma, 103*(1–2), 51–77. https://doi.org/10.1016/S0016-7061(01)00069-6.

Kyriakidis, P. C., & Journel, A. G. (1999). Geostatistical space-time models: A review. *Mathematical Geology, 31*(6), 651–684. https://doi.org/10.1023/A:1007528426688.

Lantuéjoul, C. (2002). *Geostatistical simulations: Models and algorithms*. Berlin: Springer.

Lee, S. Y., Carle, S. F., & Fogg, G. E. (2007). Geologic heterogeneity and a comparison of two geostatistical models: Sequential Gaussian and transition probability-based geostatistical simulation. *Advances in Water Resources, 30*(9), 1914–1932. https://doi.org/10.1016/j.advwatres.2007.03.005.

Mackay, J. D., Jackson, C. R., Brookshaw, A., Scaife, A. A., Cook, J., & Ward, R. S. (2015). Seasonal forecasting of groundwater levels in principal aquifers of the United Kingdom. *Journal of Hydrology, 530,* 815–828. https://doi.org/10.1016/j.jhydrol.2015.10.018.

Manzione, R. L., Knotters, M., Heuvelink, G. B. M., Von Asmuth, J. R., & Camara, G. (2008). Predictive risk mapping of water table depths in a Brazilian Cerrado area. In A. Stein, S. Wenzhong, & W. Bijker (Eds.), *Quality aspects in spatial data mining* (pp. 73–89). Boca Raton: CRC.

Manzione, R. L., Soldera, B. C., & Wendland, E. C. (2017). Groundwater system response at sites with different agricultural land uses: Case of the Guarani Aquifer outcrop area, Brotas/SP-Brazil. *Hydrological Sciences Journal, 62*(1), 28–35. https://doi.org/10.1080/02626667.2016.1154148.

Melo, A. C. G. & Durigan, G. (2011). *Plano de manejo da Estação Ecológica de Santa Bárbara, Secretaria do Meio Ambiente do Governo do Estado de São Paulo (Instituto Florestal)*. http://iflorestal.sp.gov.br/institutoflorestal/files/2013/03/Plano_de_Manejo_EEc_Santa_Barbara.pdf. Accessed July 13, 2017.

Mohanty, S., Jha, M. K., Raul, S. K., Panda, R. K., & Sudheer, K. P. (2015). Using artificial neural network approach for simultaneous forecasting of weekly groundwater levels at multiple sites. *Water Resources Management, 29*(15), 5521–5532. https://doi.org/10.1007/s11269-015-1132-6.

Montgomery, D. C., Jennings, C. L., & Kulahci, M. (2015). *Introduction to time series analysis and forecasting* (2nd ed.). Hoboken: Wiley.

Newbold, P., Agiakloglou, C., & Miller, J. (1993). Long-term inference based on short-term forecasting models. In T. S. Rao (Ed.), *Developments in time series analysis* (pp. 9–25). London: Chapman & Hall.

Nourani, V., Ejlali, R. G., & Alami, M. T. (2011). Spatiotemporal groundwater level forecasting in coastal aquifers by hybrid neural network-geostatistics model: A case study. *Environmental Engineering Science, 28*(3), 217–228. https://doi.org/10.1089/ees.2010.0174.

Nunes, R., & Almeida, J. A. (2010). Parallelization of sequential Gaussian, indicator and direct simulation algorithms. *Computers & Geosciences, 36*(8), 1042–1052. https://doi.org/10.1016/j.cageo.2010.03.005.

Olea, R. A. (1999). *Geostatistics for engineers and earth scientists*. New York: Springer.

Omran, E. E. (2016). A stochastic simulation model to early predict susceptible areas to water table level fluctuations in North Sinai, Egypt. *The Egyptian Journal of Remote Sensing and Space Science, 19*(2), 235–257. https://doi.org/10.1016/j.ejrs.2016.03.001.

Peterson, T. J., & Western, A. W. (2014). Nonlinear time-series modeling of unconfined groundwater head. *Water Resources Research, 50,* 8330–8355. https://doi.org/10.1002/2013WR014800.

Prado, R., & West, M. (2010). *Time series: Modeling, computation, and inference*. Boca Raton: CRC Press.

R Core Team. (2016). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org/. Accessed July 12, 2017.

Remy, N., Boucher, A., & Wu, J. (2011). *Applied geostatistics with SGeMS—A user's guide*. Cambridge: Cambridge University Press.

Renard, P. (2007). Stochastic hydrogeology: What professionals really need? *Ground Water, 45*(5), 531–541. https://doi.org/10.1111/j.1745-6584.2007.00340.

Santos, T., & Bonotto, D. (2014). ^{222}Rn, ^{226}Ra and hydrochemistry in the Bauru Aquifer System, São José do Rio Preto (SP), Brazil. *Applied Radiation and Isotopes, 86,* 109–117. https://doi.org/10.1016/j.apradiso.2013.12.003.

Schwarz, G. (1978). Estimating the dimension of a model. *The Annals of Statistics, 6*(2), 461–464. http://www.jstor.org/stable/2958889. Accessed Juy 13, 2017.

Shumway, R. H., & Stoffer, D. S. (2011). *Time series analysis and its application: With R examples* (3rd ed.). New York: Springer.

Singh, C. K., & Katpatal, Y. B. (2017). A GIS-based design of groundwater level monitoring network using multicriteria analysis and geostatistical method. *Water Resources Management, 31*(13), 4149–4163. https://doi.org/10.1007/s11269-017-1737-z.

Sun, Y., Wendi, D., Kim, D. E., & Liong, S.-Y. (2016). Technical note: Application of artificial neural networks in groundwater table forecasting—A case study in a Singapore swamp forest. *Hydrology and Earth System Sciences, 20,* 1405–1412. https://doi.org/10.5194/hess-20-1405-2016.

Tsanis, I. K., Coulibaly, P., & Daliakopoulos, I. N. (2008). Improving groundwater level forecasting with a feedforward neural network and linearly regressed projected precipitation. *Journal of Hydroinformatics, 10*(4), 317–330. https://doi.org/10.2166/hydro.2008.006.

Varouchakis, E. A. (2016). Modeling of temporal groundwater level variations based on a Kalman filter adaptation algorithm with exogenous inputs. *Journal of Hydroinformatics, 19*(3), 191–206. https://doi.org/10.2166/hydro.2016.063.

Von Asmuth, J. R., Maas, K., Bakker, M., & Petersen, J. (2008). Modeling time series of ground water head fluctuation subjected to multiple stresses. *Ground Water, 46*(1), 30–40. https://doi.org/10.1111/j.1745-6584.2007.00382.x.

Woodward, W. A., Gray, H. L., & Elliott, A. C. (2011). *Applied time series analysis*. Boca Raton: CRC Press.

Wunsch, A., Liesch, T., & Broda, S. (2018). Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX). *Journal of Hydrology*. https://doi.org/10.1016/j.jhydrol.2018.01.045. **(in press)**.

Yi, M.-J., & Lee, K.-K. (2004). Transfer function-noise modelling of irregularly observed groundwater heads using precipitation data. *Journal of Hydrology, 288,* 272–287. https://doi.org/10.1016/j.jhydrol.2003.10.020.

Yihdego, Y., & Webb, J. A. (2011). Modeling of bore hydrographs to determine the impact of climate and land-use change in a temperate subhumid region of southeastern Australia. *Hydrogelogy Journal, 19,* 877–887. https://doi.org/10.1007/s10040-011-0726-1.

Zhang, Y. K., & Zhang, D. (2004). Forum: The state of stochastic hydrology. *Stochastic Environment Research and Risk Assessment, 18*(4), 265–265. https://doi.org/10.1007/s00477-004-0190-8.