Science of the Total Environment 881 (2023) 163328

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

A parsimonious methodological framework for short-term forecasting of groundwater levels

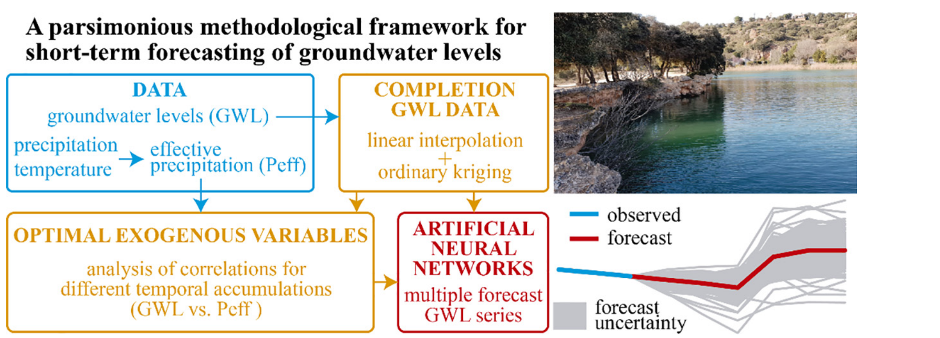
A.J. Collados-Lara [a,*], D. Pulido-Velazquez [b], L.G.B. Ruiz [c], M.C. Pegalajar [d], E. Pardo-Igúzquiza [e], L. Baena-Ruiz [b]

[a] *Department of Civil Engineering, University of Granada, Water Institute, 18003 Granada, Spain*
[b] *Spanish Geological Survey (IGME-CSIC), 18006 Granada, Spain*
[c] *Department of Software Engineering, University of Granada, 18014 Granada, Spain*
[d] *Department of Computer Science and Artificial Intelligence, University of Granada, 18014 Granada, Spain*
[e] *Spanish Geological Survey (IGME-CSIC), 28003 Madrid, Spain*

HIGHLIGHTS

- Useful to estimate levels of drought alert and anticipate management responses
- It integrates artificial neural networks, geostatistics and climatic variables.
- NAR is the best approach in well locations with lower R^2 between GWL and P and Pef.
- NARX and Elman approaches were selected >70 % of the instances.
- We obtained a mean RMSE of 0.90 m forecasting test for the 51 wells studied.

GRAPHICAL ABSTRACT



ARTICLE INFO

Editor: Christian Herrera

Keywords:
Groundwater levels
Ordinary kriging
Effective precipitation
Artificial neural networks

ABSTRACT

Groundwater plays a significant role as a strategic resource in reducing the impact of droughts. In spite of its importance, there are still many groundwater bodies in which there is not enough monitoring data to define classic distributed mathematical models to forecast future potential levels. The main aim of this study is to propose and evaluate a novel parsimonious integrated method for the short-term forecasting of groundwater levels. It has low requirements in term of data, and it is operational and relatively easy to apply. It uses geostatistics, optimal meteorological exogenous variables and artificial neural networks. We have illustrated our method in the aquifer "Campo de Montiel" (Spain). The analysis of optimal exogenous variables revealed that, in general, the wells with stronger correlations with precipitation are located closer to the central part of the aquifer. NAR, which does not consider secondary information, is the best approach for 25.5 % of the cases and is associated with well locations with lower R^2 between groundwater levels and precipitation. Amongst the approaches with exogenous variables, the ones that use effective precipitation have been selected more times as the best experiments. NARX and Elman using effective precipitation had the best approaches with 21.6 % and 29.4 % of the cases respectively. For the selected approaches, we obtained a mean RMSE of 1.14 m in the test and 0.76, 0.92, 0.92, 0.87, 0.90, and 1.05 m for the forecasting test for months 1 to 6 respectively for the 51 wells, but the accuracy of the results can vary depending on the well. The interquartile range of the RMSE is around 2 m for the test and forecasting test. The uncertainty of the forecasting is also considered by generating multiple groundwater level series.

* Corresponding author.
E-mail address: ajcollados@ugr.es (A.J. Collados-Lara).

http://dx.doi.org/10.1016/j.scitotenv.2023.163328
Received 31 January 2023; Received in revised form 14 March 2023; Accepted 2 April 2023
Available online 6 April 2023

1. Introduction

One of the main challenges of water resource management in areas with scarce resources is how to deal with droughts and their propagation (Barker et al., 2016; Hidalgo-Hidalgo et al., 2022). These problems may become exacerbated in the future due to climate change (Collados-Lara et al., 2020, 2022), especially in the Mediterranean basin (Cramer et al., 2018; Tramblay et al., 2020) and areas with conflicts between wetland conservation and intensive agriculture (Ahmadi et al., 2021; Pulido-Velazquez et al., 2023). In these systems, groundwater may play a significant role as a strategic resource to reduce the impact of droughts (Leduc et al., 2017; Pulido-Velazquez et al., 2020).

In order to analyse the potential management strategies in a rational way, in addition to performing continuous monitoring activities and methods to estimate the hydrodynamic parameters in the aquifer (Dashti et al., 2023), we need to use models which help to understand the system behaviour better and to forecast the potential future status of the system (Linés et al., 2018; Gomez-Gomez et al., 2022). The models are used to assess the potential future long horizon impact, including the propagation of climate change effects on aquifer recharge (Pardo-Igúzquiza et al., 2019; Dubois et al., 2022) and groundwater levels (Moseki, 2018; Pulido-Velazquez et al., 2018), but also for short-term future forecasts of groundwater recharge or status (Zaadnoordijk et al., 2019). These short-term predictions can even be applied to estimate levels of drought alert in accordance with the historical and forecasted groundwater levels (Pham et al., 2022; Mackay et al., 2015).

In this context, machine learning (ML) models are widely used for predicting groundwater levels and dynamics, as they can capture complex non-linear relationships between input and output variables. ML is a subset of AI that involves the use of algorithms and statistical models to enable machines to learn from data without being explicitly programmed. Artificial intelligence (AI) refers to the simulation of human intelligence in machines that are programmed to perform tasks that would usually require human intelligence, such as decision-making.

Each ML model has strengths and limitations, and their performance may depend on various factors such as data quality, model structure, and parameter selection, amongst others. Many studies can be found in this regard. In (Chen et al., 2020), the authors compare one numerical model and three ML algorithms such as multi-layer perceptron, radial basis function network and extreme machine learning to simulate the groundwater dynamics of the middle reaches in China. The authors of (Rohde et al., 2021) used satellite-based remote sensing data to predict groundwater levels under different ecosystems across California. They used random forests to relate groundwater levels to vegetation indices. Finally, a systematic review can be found in (Tao et al., 2022) of some studies that used ML models for this purpose.

In spite of their utility, there are still many groundwater bodies in which there is not enough monitoring and/or there are no mathematical models to forecast future potential levels (Mengistu et al., 2021). In recent years, there has been an increased application of ML approaches (Jimeno-Sáez et al., 2017; Martinsen et al., 2022) in addition to the classical stochastic models, such as auto-regressive models, in the hydrology field for basin (Senent-Aparicio et al., 2018; Pulido-Velazquez et al., 2008) and aquifer scale (Llopis-Albert and Pulido-Velazquez, 2015; Baena-Ruiz et al., 2018) modelling. However, errors in the structure of the model, parameters, and data can lead to both random and systematic errors in the output of a calibrated model (Xu et al., 2014). Therefore, there is a need for improved modelling techniques that can better account for these uncertainties and provide accurate forecasts of groundwater levels.

Recent advances in ML techniques have shown promise in the field of hydrology as an alternative to traditional numerical models for forecasting future groundwater levels (Tao et al., 2022; Aguilera et al., 2019). One advantage of the ML approach is that it requires less data and time for model definition and calibration, compared to the classical distributed models (Tao et al., 2022; Aguilera et al., 2019). In these lines, an artificial neural network (ANN) is a type of ML algorithm that is inspired by the structure and function of the human brain. It consists of interconnected nodes, or neurons, which are organised into layers that process information and can learn to recognize patterns in data, making it a powerful tool for hydrology applications. Coupling wavelet analysis with ANNs has also been explored as a way to enhance the accuracy of the predictions (Adamowski and Chan, 2011). ML is also combined with other methods to model the spatio-temporal dynamic (Nourani and Mousavi, 2016) to improve the accuracy of the simulations. ANNs can be used as a surrogate model in hydrological processes (Zhang et al., 2020). Surrogate modelling is a technique commonly used in engineering and scientific applications to model complex systems when direct measurement or computation of the outcome is difficult or costly (Razavi et al., 2012). The surrogate model serves as a proxy for the original system and can be used to make predictions or optimize the system with a reduced computational burden. The surrogate model is built based on a limited set of input/output data pairs from the original system and can be constructed using various methods (Zhang et al., 2020). Similarly, response surface modelling (Kay et al., 2021) is a mathematical and statistical technique useful for developing and improving processes. It is used to analyse the complex relationships between multiple independent variables and their effect on the variable of interest. It may help to identify the optimal values of the independent variables that maximise or minimise the response variable. This technique is particularly useful when multiple input factors are expected to have an impact on the output.

ML approaches have been used to consider different exogenous variables, such as climate variables, which can improve the accuracy of the forecast (Hussein et al., 2020).The most commonly used exogenous variables for hydrological modelling are precipitation and temperature, but in some cases, combined variables such as effective precipitation have also been used (Tao et al., 2022). Exogenous variables may also help to understand the variability of groundwater quantity and quality in the same way as groundwater clustering techniques (Nourani et al., 2022). However, these exogenous or target variables may require additional processing using geostatistical techniques to convert them into useful information for the proposed forecast approach (Collados-Lara et al., 2018; Bhat et al., 2015). Most of the models aim at assessing the potential impact of droughts for management purposes operate at a monthly scale (Mukherjee et al., 2018). This is because extreme events such as droughts are not characterized by a fast response of the system, unlike flood analyses that require consideration at daily or even smaller temporal resolutions to simulate potential impacts (Didovets et al., 2019). Analyses of droughts and their propagation are usually performed by aggregating monthly results into seasonal, yearly (the most common approach), or longer periods.

In this study, we propose a novel methodological framework for short-term forecasting of groundwater levels (GWL) on a monthly scale by using meteorological data, geostatistics and artificial neural networks (ANN). Generally, for a specific problem, multiple ANN structures can satisfactorily represent the training data, and for each structure, there are many appropriate sets of network weights and biases (Razavi et al., 2012). In this study, we have tested different ANN approaches, exogenous variables, number of neurons and delays.

The aim of this study is to propose and analyse the potential of a novel parsimonious integrated method for short-term forecasting of groundwater levels. It has low requirements in terms of data, and it is operational and relatively easy to apply. The main novelty of our method with respect to previous research studies is that it integrates (1) an ordinary kriging procedure to complete gaps in GWL series, (2) a novel approach to select optimal exogenous variables based on GWL response time to meteorological forcing, (3) a detailed analyses of several ANN approaches and exogenous variables within each well location has also been performed, and (4) it generates multiple simulations of future GWL series in order to take into account uncertainty within forecast systems. The proposed methodological framework is illustrated through the analysis of the case study of the Campo de Montiel groundwater body (GWB), which is located in the Upper Guadiana basin (south-eastern Spain). This area highlights a conflict between groundwater-dependent ecosystems and groundwater pumping to supply irrigation demands and the proposed approach could be very useful to estimate levels of drought alert and anticipate the management responses. However, our method is a parsimonious approach (from the point of view of the needed data) which can easily be applied to any case study.

2. Methodology

The methodological framework proposed (Fig. 1) requires (1) the completion of the GWL data, (2) the analysis of the correlations of GWL data with precipitation and effective precipitation to obtain the optimal exogenous variable, and (3) the application of the ANN approach proposed. These methodologies are explained in the following sub-sections.

2.1. GWL data completion

We applied two completion procedures, a temporal linear interpolation procedure applied to each well location and a geostatistical spatio-temporal interpolation applied to the whole well location. The linear interpolation was applied to the data gaps lower than or equal to 6 months. The geostatistical interpolation was applied to the remaining gaps in the GWL data. GWL is a variable with spatial continuity which changes gradually. The ordinary kriging (OK) approach takes into account the spatial correlation information to estimate fields (Collados-Lara et al., 2018). OK uses data of the target variable to obtain the estimates and the variogram to quantify the spatial correlation. In our case study, we fitted a Gaussian variogram to the mean GWL data. We tested two approaches based on OK, applied to the GWL data and applied to the difference between the GWL data and the average GWL data for each well location. Both procedures were compared through a cross-validation procedure. In geostatistics the cross-validation (Chilès and Delfiner, 1999) is generally accepted as the following the procedure which is also known as the leave-one-out cross-validation: 0) The variogram of the variable of interest (GWL in our case) is estimated using the complete set of n experimental data, 1) one datum is eliminated from the data set, 2) the rest of the n-1 data are used to produce an estimate of the target variable at the location where the datum was eliminated, 3) the true error incurred in this process is calculated by the difference between the actual value and the estimated value). 4) Steps 1 to 3 are repeated for the n experimental data. 5) Cross-validation statistics are calculated by using the n true errors. In this study we used two statistics: the mean error (ME) and the root mean squared error (RMSE). The ME is the bias of the estimation, where the value should be around zero. The RMSE is the accuracy of the estimate and the value should be as small as possible.

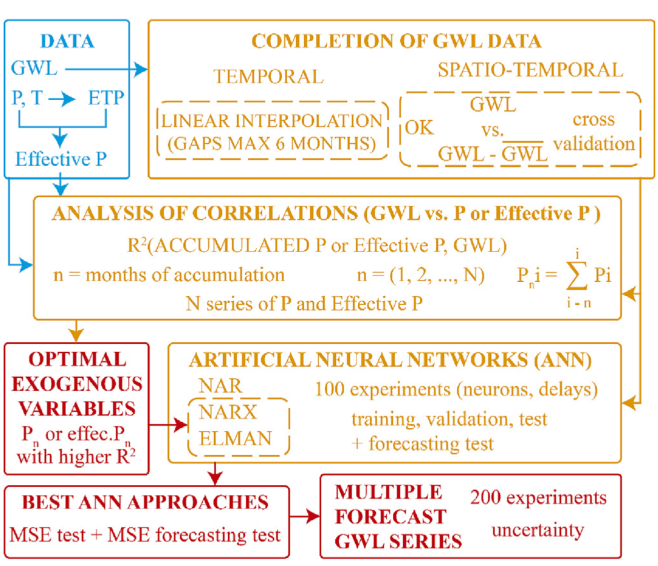


Fig. 1. Flow chart of the proposed method.

2.2. Analysis of the correlation of GWL with exogenous variables

The correlation of GWL vs. precipitation (P) and effective precipitation (Pef) were analysed for all the well locations to obtain the exogenous variables that can explain in part the GWL dynamics. Pef was calculated using P and potential evapotranspiration on a daily scale. Potential evapotranspiration was calculated from the temperatures by using the Hargraves method (Hargreaves and Allen, 2003). We accumulated P and Pef considering different backwards windows of time (see Fig. 1). Thus, if we consider accumulation periods up to the last 24 months we have 24 time series of P and Pef. The correlation coefficient (R^2) of GWL vs. P and Pef was calculated for each well location by using the different accumulated series. Note that depending on the well location the relationship with the studied exogenous variables maybe different. The GWL can be explained in part by P or Pef if we obtain high R^2 or less clearly if we obtain low R^2. The response of GWL to the changes of P or Pef can be fast if we are able to obtain higher correlations for lower accumulation periods or slow if we obtain higher correlations for higher accumulation periods.

2.3. ANN application

ML models are widely used in hydrological modelling to learn the statistical correlation between random variables of interest, such as rainfall, runoff, water level, water quality, etc. (Yang and Chui, 2021). Some of the common ML models that have been applied in this field are ANN, support vector machines (SVM), random forests (RF) k-nearest neighbours (kNN) amongst others (Herath et al., 2021; Shen et al., 2021). These models have different advantages and disadvantages in terms of accuracy, complexity and interpretability. In this context, ANNs are often more suitable for non-linear and dynamic problems (Essenfelder and Giupponi, 2020) though they may also suffer from overfitting or lack of transparency. Another ANN-based technique is deep learning widely used to learn complex patterns from large amounts of data (Shen and Lawson, 2021). However, we did not select deep learning for our problem because it requires more computational resources, data availability and domain knowledge than other models. Thus, we selected three types of ANNs: NAR, NARX and Elman neural networks. These models can capture the temporal dependencies and feedback in hydrological systems using recurrent connections or exogenous inputs. We applied the three types of ANNs for each well location. In the cases of NARX and Elman we tested the time series of P and Pef (Section 2.2) as exogenous variables. The ANN approaches were applied by using the functions *narnet*, *narxnet* and *elmannet* available in Matlab®. In addition to the traditional procedure for modelling with ANN (training, validation, and test) we included a forecasting test (see Fig. 2). Note that in training, for the validation and test the estimated value is obtained by using the previously available data and the objective of this paper is to obtain a forecast of 6 months. In the forecasting test we performed 7 forecasts of 6 months by using the last 12 months of the GWL series (Fig. 2). The rest of the time series of GWL (N-12 months) were used to perform the training, validation and test experiments (we used a random procedure to split the N-12 months into three parts). The algorithm Levenberg-Marquardt was used for the training of the ANN. The statistic of goodness of fit used for the training, validation, test, and forecasting test was the mean squared error (MSE). The application of the ANN was done by using normalized series (from 0 to 1) of GWL and exogenous variables.

For each well location and ANN approach (5 in total, NAR, NARX P, NARX Pef, Elman P, Elman Pef) we performed 100 experiments varying the delay ($1-10$) and neurons (1–10). We repeated each experiment 20 times in order to obtain a representative MSE. Note that the modelling of ANN includes random procedures that can vary the MSE in each execution using the same configuration. The selection of the best experiment for each well location and ANN approach was done by considering the mean value of MSE in the test and forecasting test. The configuration of the best experiments was used to model GWL from data 1 to N-12 and to perform 7 forecasts by using the last 12 months (see Fig. 2). We repeated the best

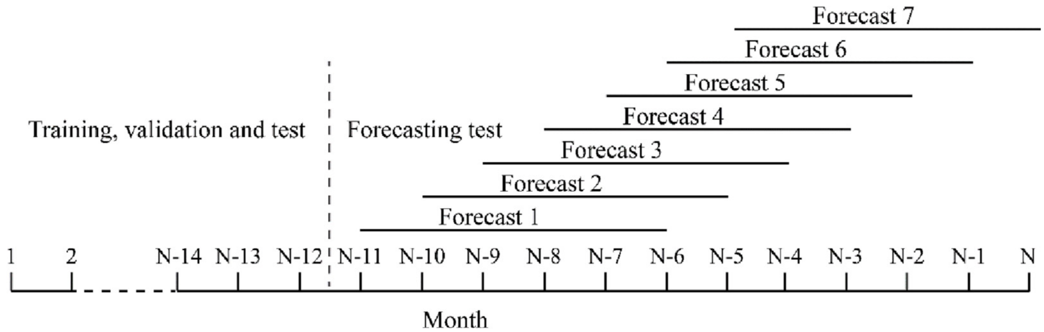


Fig. 2. Dataset use scheme in the ANN application.

experiments 200 times to obtain a range of uncertainty associated with the modelling. To account for these inherent uncertainties in the modelling process, we carried out a simulation approach. We generated multiple simulations of GWL series and predictions based on different input data sets. This allowed us to capture the range of potential outcomes and estimate the uncertainty associated with each forecast. To assess the consistency of our predictions across different forecast horizons, we conducted sensitivity analyses. By examining the prediction intervals of our simulations, we were able to gain insights into the sources and magnitudes of uncertainty associated with the forecasting process.

3. Case study and data

The Campo de Montiel groundwater body (GWB) is located in the Upper Guadiana basin (south-eastern Spain) (see Fig. 3). The elevation varies from 657 to 1098 m a.s.l. in the GWB and it has a surface of around 2220 km^2. In this GWB there is a strong natural interaction between the groundwater and the surface water. Note that the Lagunas de Ruidera wetland (Natural Park and a Ramsar area) and Peñarroya reservoir are located in this area (Collados-Lara et al., 2021). This area highlights a conflict between groundwater-dependent ecosystems and groundwater pumping to supply demands (mainly irrigation demands). The site is critical for the functioning of the regional hydrological system and serves as an important water reservoir in this semi-arid region. The mean temperature in the GWB in the period 1950–2015 was14.2 °C and the mean minimum and maximum temperatures were 7.9 °C and 20.5 °C respectively. The mean annual precipitation is 503 mm/year and the effective precipitation is 331.0 mm/year. The main geological formations are limestone and dolomite and the GWB behaves like a free aquifer. The recharge of the GWB occurs only through direct infiltration of the rainfall, whilst the discharge is carried out fundamentally through the drainage of the GWB towards the surface water courses, resulting in the source of the Guadiana Alto-Pinilla, Azuer (and Cañamares) and Córcoles rivers. Furthermore, another part of the discharge of the GWB is the groundwater flow to the GWB Mancha Occidental I and II, and a third part feeds the wetlands.

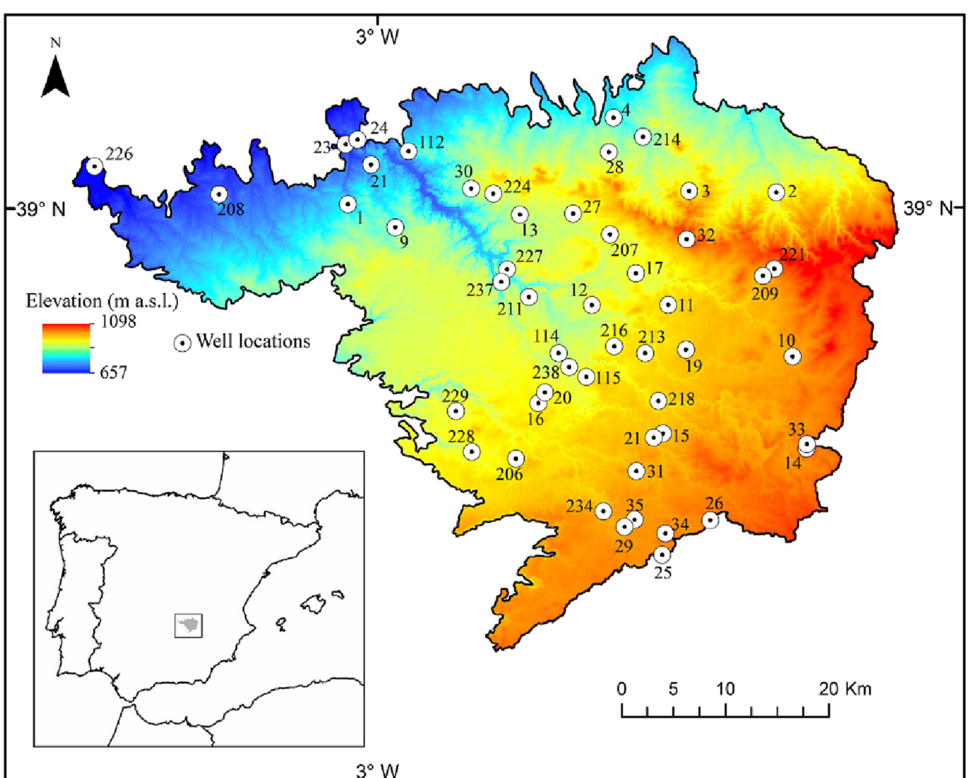


Fig. 3. Location of the case study, well locations, and elevation map.

The climatological variables, precipitation and temperature (mean, maximum and minimum) were obtained from the Spain02 (v5) project (Herrera et al., 2016; Kotlarski et al., 2019). The groundwater levels (GWL) data in the 51 well locations (Fig. 3) were provided by the Guadiana Hydrographic Confederation (https://www.chguadiana.es/). We considered that the availability of GWL data before January 2000 was not enough to apply our method. Therefore we considered as a study time period from January 2000 to January 2020 (252 months). The available monthly GWL data in this period (Fig. 4) is different for each well location and require additional geostatistical processing techniques to obtain complete series.

4. Results

4.1. GWL data completion and exogenous variables

The OK was applied to the GWL and GWL-$\overline{\text{GWL}}$ (see Section 2.1) to complete the data. The second procedure shows the best results in the cross-validation experiment in terms of ME and RMSE. The ME applying OK to GWL is 0.77 m whilst the RMSE when OK is applied to GWL-$\overline{\text{GWL}}$ is 0.02 m (Fig. 5). In the case of RMSE these values are respectively 27.9 and 2.9 m. Therefore, the second procedure was used to complete the GWL series. The completed GWL series (Fig. 6a–d) vary from 620 to 1015 m a.s.l. The Pef series (calculated as explained in Section 2.2) are represented with P series in Fig. 6e. For the case study the Pef represent 65.8 % of P.

4.2. Analysis of the correlation of GWL with exogenous variables

The R^2 of the GWL with P varies from 0.13 to 0.73 in the 51 well locations, the mean value is 0.47 and the median is 0.53 (Fig. 7a). The accumulation period for which this R^2 is reached varies from 3 to 30 months, the mean value is 14.3 and the median 16 (Fig. 7a). The R^2 of the GWL with Pef varies from 0.13 to 0.81 in the 51 well locations, the mean value is 0.50 and the median is 0.58 (Fig. 7b). The accumulation period for which this R^2 is reached varies from 3 to 30 months, the mean value is 14.4 and the median 12 (Fig. 7b). In the case of P 82.4 % of the well locations show a moderate correlation ($r > 0.5$) and a 41.2 % show a strong correlation ($r > 0.75$). The values in the case of Pef are higher, 88.2 % and 52.9 % respectively. The spatial distribution of the well locations according to their correlation with Pef is shown in Fig. 8. In general the wells with stronger correlations are located closer to the central valley of the GWB (see Fig. 3) and the wells with weak correlations are located close to the GWB boundaries.

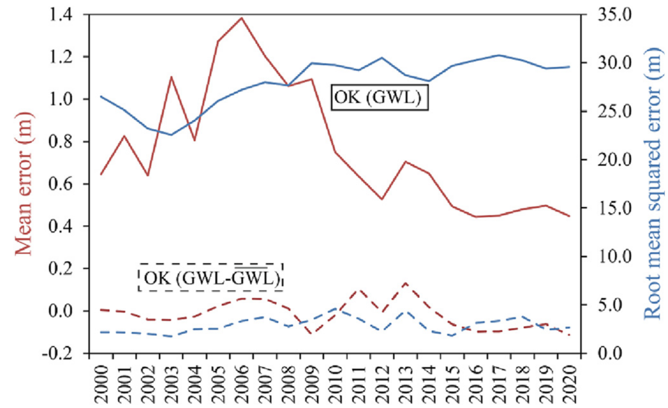


Fig. 5. Mean error and root mean squared error obtained in the cross validation experiment for the OK applied to GWL and OK applied to the difference of GWL and mean GWL.

4.3. The application of artificial neural networks

In order to test the performance of the proposed models, we used a systematic testing approach to determine the optimal ANN parameters. Specifically, we focused on two key parameters, the number of neurons in the hidden layer and the delay parameter, which controls the number of lagged inputs to the ANN. We used a trial-and-error approach to test various numbers of neurons and delays and evaluated their performance using statistical metrics such as RMSE, MSE and R^2. Although other ANN parameters, such as learning rate and momentum, can also influence model learning, we were unable to include them in our analysis due to length restrictions. The application of the ANN was performed by using normalized series (from 0 to 1) of GWL, P and Pef but the final results are presented as denormalized. As explained in Section 2.3, for each well location and ANN approach (5 in total, NAR, NARX P, NARX Pef, Elman P, Elman Pef) we performed 100 experiments varying the delay (1–10) and neurons

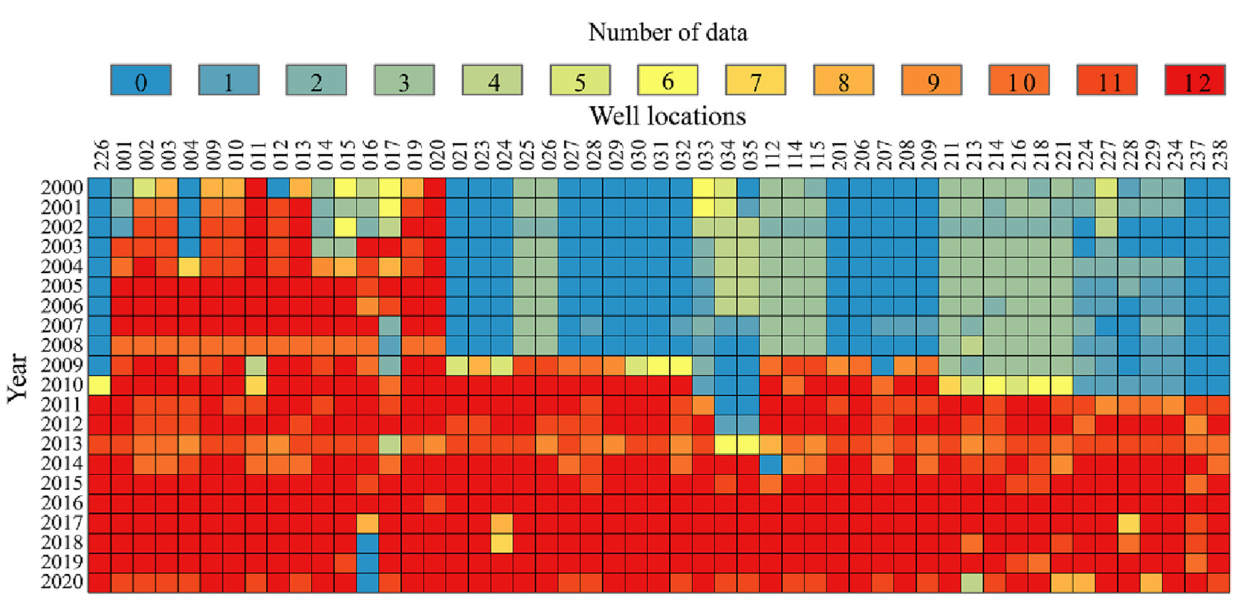


Fig. 4. GWL data availability in each well location.

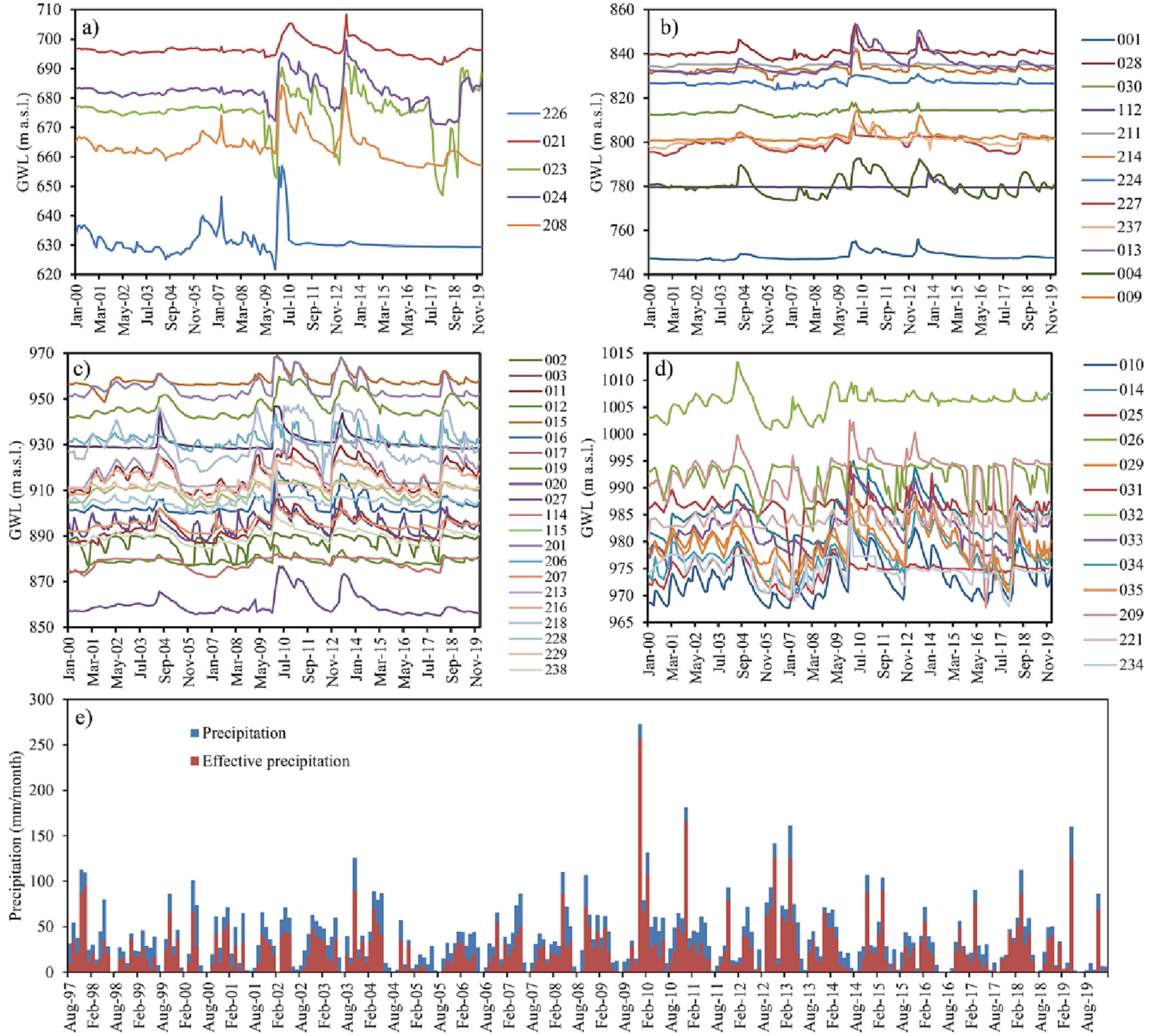


Fig. 6. Completed groundwater levels series for the studied period (2000–2020) (a) series with mean value between 620 and 700 m a.s.l, (b) series with mean value between 740 and 840 m a.s.l, (c) series with mean value between 860 and 960 m a.s.l, (d) series with mean value between 970 and 1005 m a.s.l, and precipitation and effective precipitation for the GWB in the period (1997–2020) (e).

(1–10). Note that the proposed methodology was applied for each well location. However we also analysed the patterns of the whole case study by studying mean values of MSE. The mean MSE for the test and forecasting test for the different months of the forecast (from 1 to 6 months) of each experiment and ANN approach is shown in Fig. 9. In general the MSE of the test and the forecast of month 1 are similar or even lower for the forecast and it increases with the months of forecast for all the experiments and ANN approaches. In the case of NAR the lowest MSE is obtained for the experiment with delay = 1 and neurons = 1 and it increases with neurons (Fig.10a) and delays. The same pattern is observed for NARX P (Fig. 10c and d) and NARX Pef (Fig. 10e and f). In the case of Elman P and Elman Pef the MSE also increases with the neurons (Fig. 10g and i respectively) but a pattern is not observed with respect to the number of delays (Fig. 10h and j).

As we pointed out before, all the ANN approaches were applied for each well location. The results of the best ANN approach for each well location are summarized in Fig. 11 and Table 1. The ANN approach that shows the lowest mean RMSE in the test is NARX Pef (Fig. 12a). However NAR shows the lowest mean RMSE for the forecasting test (Fig. 12a). NAR, which does not consider secondary information from P or Pef, is the best approach in 25.5 % of the cases and is associated with well locations with lower R^2 with P and Pef (Fig. 12a). Amongst the ANN approaches with exogenous variables, the ones that use Pef have been selected more times as the best experiments. NARX Pef and Elman Pef are the best ANN approaches with 21.6 % and 29.4 % of the cases respectively. On the other hand, NAR P and Elman P are the best ANN approaches with13.7 % and 9.8 % of the instances.

For the selected ANN approaches in each well location we obtained a mean RMSE of 1.14 m in the test and 0.76, 0.92, 0.92, 0.87, 0.90, and 1.05 m for the forecasting test for the months from 1 to 6 respectively (Fig. 12b). However the accuracy of the results can be considerably different depending on the well location (Fig. 11). Note that the interquartile

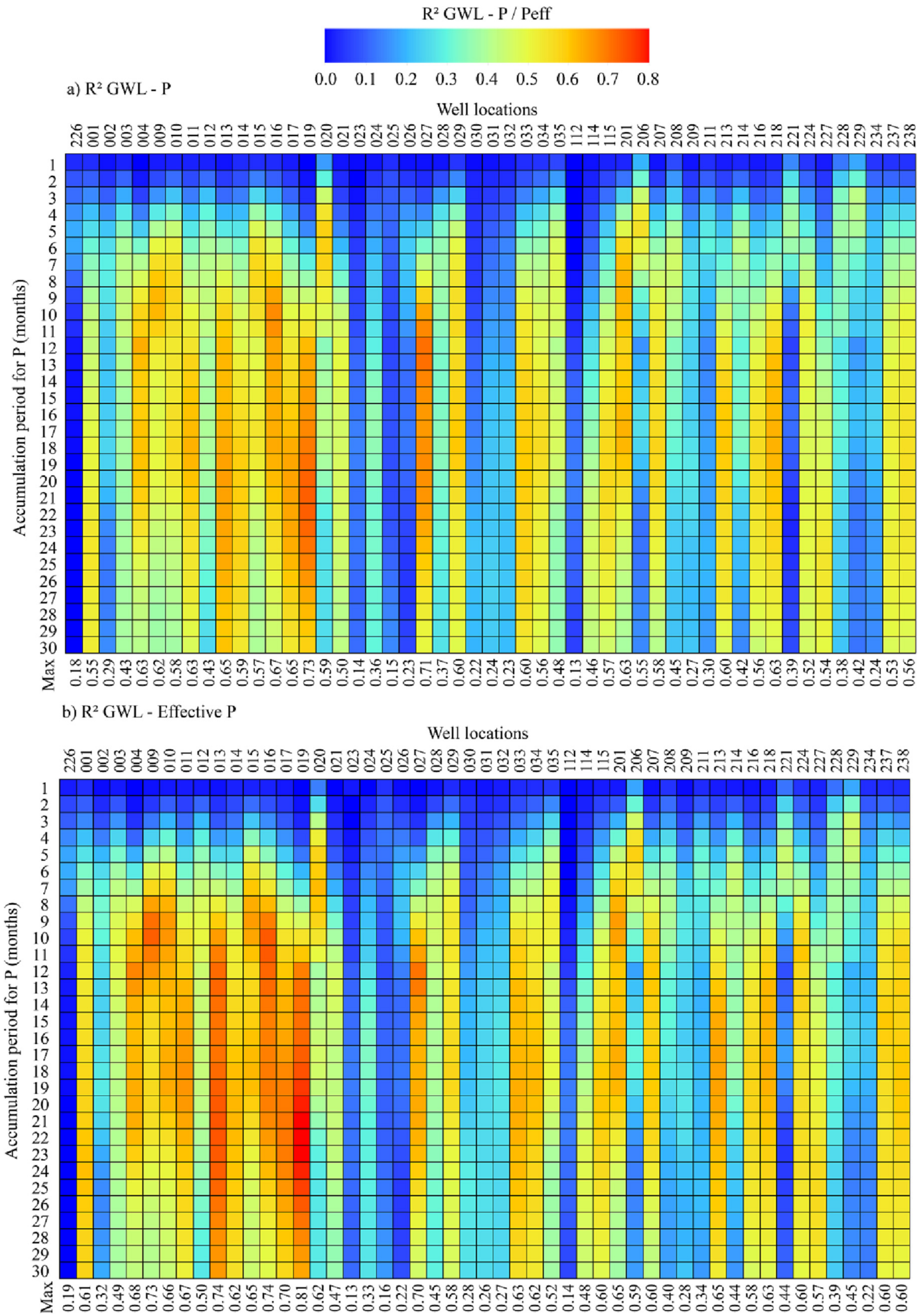


Fig. 7. Correlation coefficient obtained for the groundwater levels relationship with the tested exogenous variables: a) precipitation, b) effective precipitation for each well location.

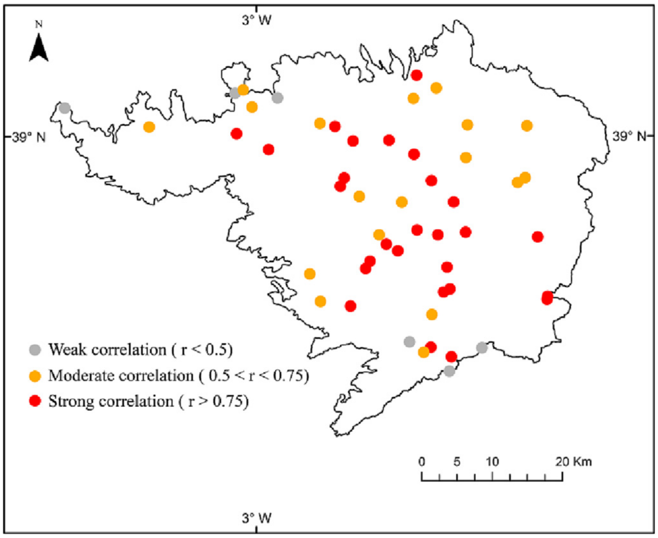


Fig. 8. Groups of well locations according their correlation with effective precipitation.

range of the RMSE is around 2 m for the test and forecasting test (Fig. 12b). In addition to the RMSE we calculated two information criteria that take into account the number of parameters used by the ANN for the selected approaches (Table 1), the Akaike information criterion (AIC) and the Bayesian information criterion (BIC). In general, ANN approaches with higher numbers of neurons and delays have a higher number of parameters and therefore the performance regarding the AIC, and BIC is worse. However, in this study the numbers of parameters were not considered to select the optimal ANN approaches.

For each well location and its corresponding best ANN approach and experiment (delay, neurons) we generated multiple estimations (a battery of 200) of the GWL series in order to take into account the modelling uncertainty. In all the cases the performance is very good. Fig. 13 shows the observed, multiple series and the average of the multiple series for four well locations. For the well location 001 (NAR approach, delay 2, neurons 2) the ensemble of the multiple series has a RMSE with respect to the observed series of 0.49 m. In the case of well 009 (NARX Pef approach, delay 3, neurons 1), well 034 (Elman P approach, delay 1, neurons 1), and well 224 (Elman Pef approach, delay 1, neurons 1) the RMSE is 0.55, 1.07, and 0.58 m respectively. We also generated multiple forecasts for all the well locations using the ANN approach and experiment (delay,

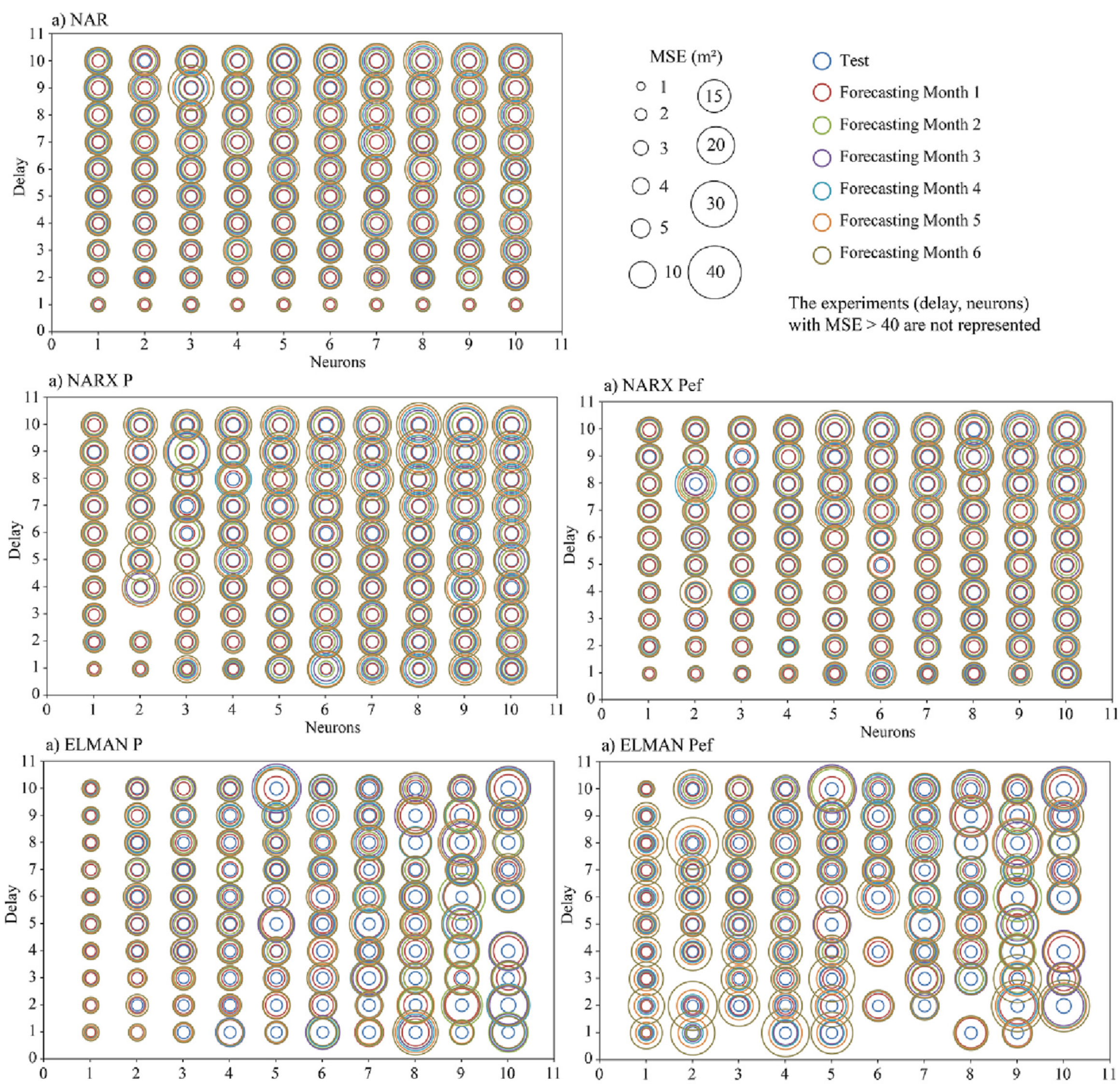


Fig. 9. Average (for all 51 well locations) MSE obtained in the different experiments (delay, number of neurons) for the test and the forecasting test (for the different months) of the ANN approaches.

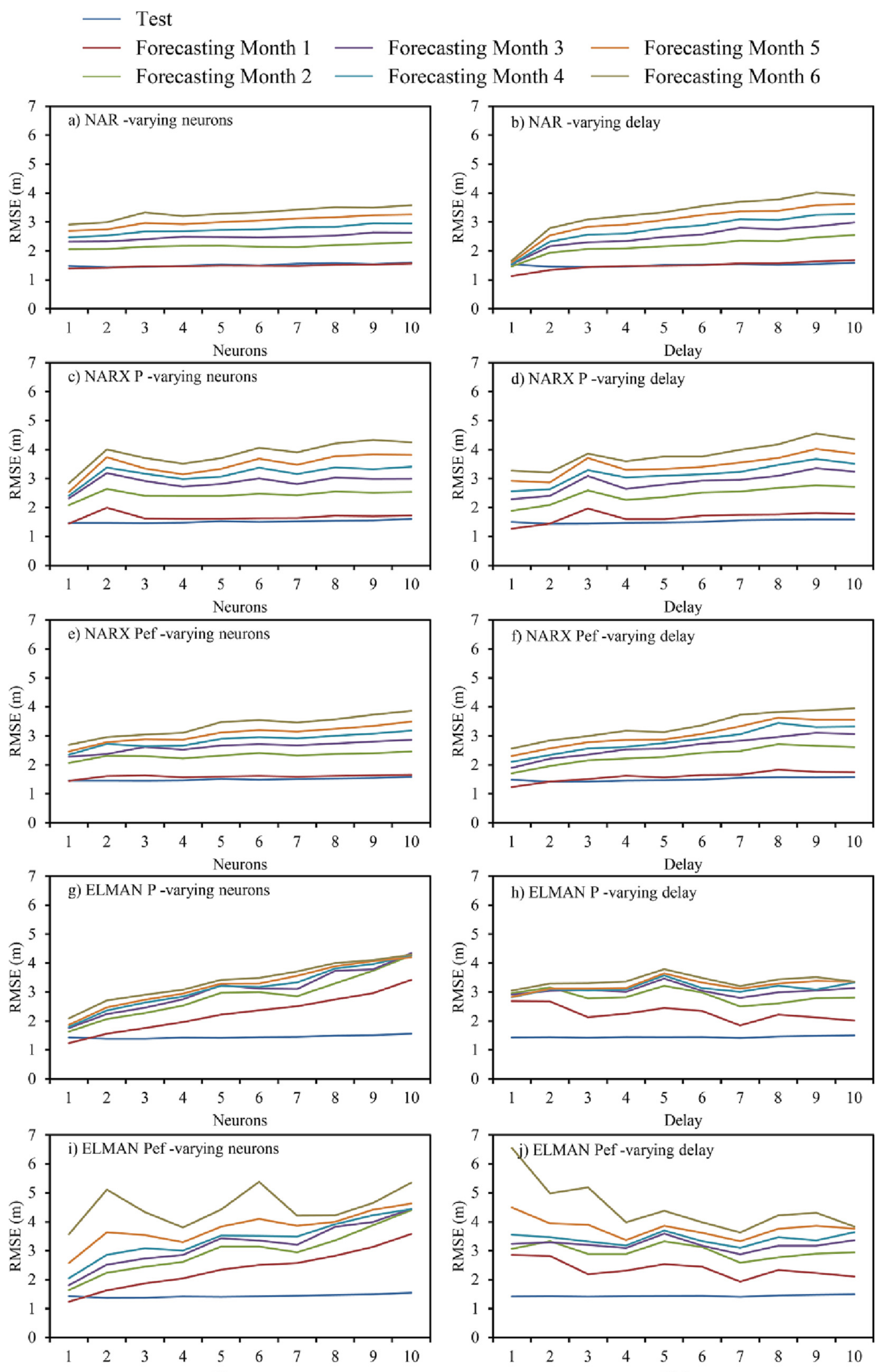


Fig. 10. Average (for all 51 well locations) RMSE obtained in the experiments varying the delay and neurons for the test and the forecasting test (for the different months) of the ANN approaches.

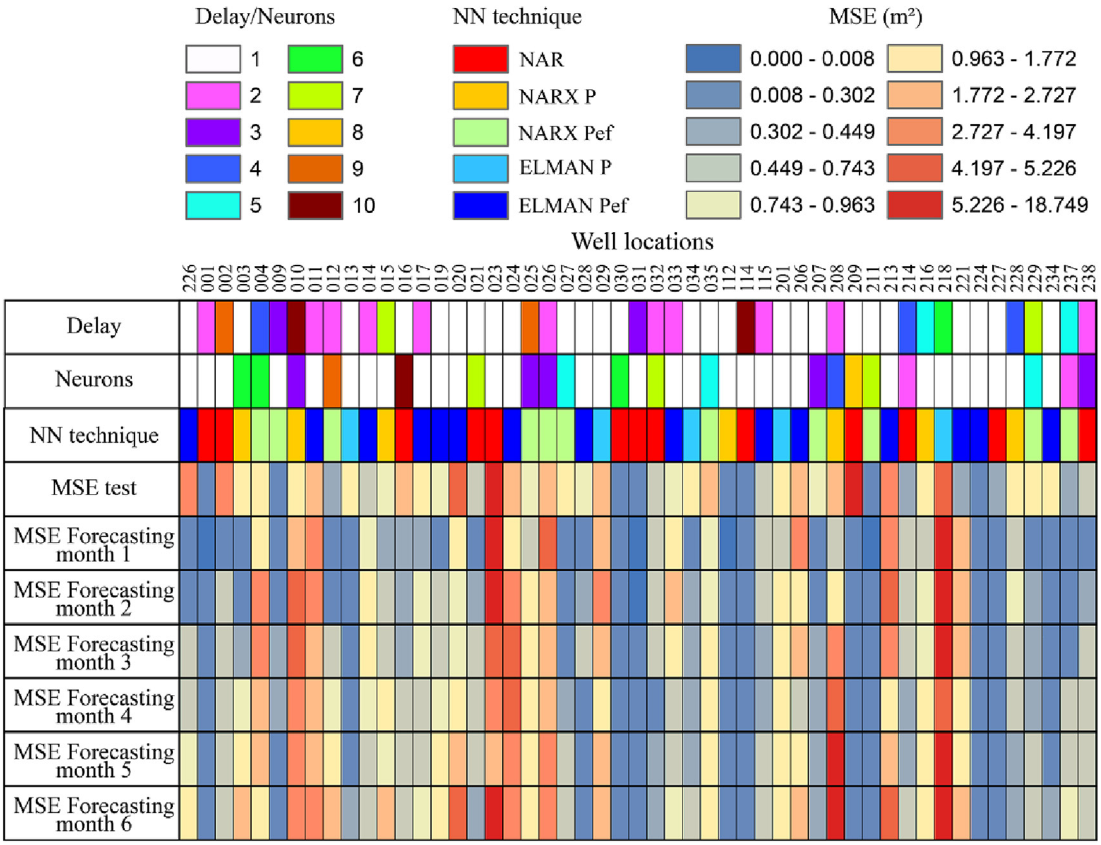


Fig. 11. The best experiments (delay, number of neurons and ANN approach) and MSE (for the test and forecasting) for each well location.

neurons) selected. We generated 7 forecasts of 6 months in length (Fig. 2). The first forecast (months 230 to 235) of four well locations is represented in Fig. 14. We represent the observed, multiple series, and the average of the multiple series for four well locations. In these cases, we show the results of well 012 (NARX Pef approach, delay 2, neurons 9), well 015 (NARX P approach, delay 7, neurons 1), well 019 (Elman Pef approach, delay 1, neurons 1), and well 228 (NARX P approach, delay 4, neurons 1). The ensemble of these multiple forecasts has a RMSE of 0.49, 0.62, 0.72, and 0.70 m respectively.

5. Discussion

We propose a novel methodological framework for short-term forecasting of GWL. It should also contribute to a better understanding of the behaviour of groundwater systems. It does not have demanding data requirements (a parsimonious approach), it is operational and relatively easy to apply. In recent years, several studies have used artificial neural networks (including exogenous precipitation data) to forecast groundwater levels (e.g. Guzman et al., 2017; Roshni et al., 2020). Few of these studies included temperature or potential evapotranspiration as exogenous variables (e.g. Di Nunno and Granata, 2020; Wunsch et al., 2021). We have investigated whether effective precipitation, which is obtained from precipitation and evapotranspiration (obtained from temperature), provides better results than precipitation. In our case study, effective precipitation was selected as an exogenous variable more times than precipitation. Note that effective precipitation is the amount of precipitation that is actually added and stored in the soil and may contribute to the groundwater resources. Moreover, as far as we know, none of these previous studies considered the possible delay between precipitation or effective precipitation and the aquifer response to select the optimal period of aggregation for the exogenous variable time series. In our case study, this delay (time to reach the maximum correlation) varies from 3 to 30 months depending on the well location. In general, the wells with stronger correlations with effective precipitation are located closer to the central valley of the GWB and the wells with weak correlations are located close to the GWB boundaries. As found in previous studies, in general considering meteorological exogenous variables improves the forecasting of GWL (Wunsch et al., 2021). In our case study the artificial neural networks that consider exogenous variables were selected for 75 % of the cases. ANN provides both, excellent prediction capability, and valuable sensitivity analyses, which can support the definition of more appropriate ground water management strategies (Coppola et al., 2005). On the other hand, incomplete data has been recognized as one of the fundamental challenges in ML (Goodfellow et al., 2016), and it is very common to have gaps in GWL series. Our methodological framework also proposes a geostatistical approach by using ordinary kriging to complete the GWL data. We found that better results are obtained in the completion if the target variable is configured as the difference between the groundwater level and the average groundwater level for each location. We also propose the generation of multiple groundwater levels simulated series to take into account the forecast uncertainty coming from the models. The results demonstrate the potential of the proposed neural network approach to simulate GWL within a historical period, where the forecast can be validated, due to the existence of real climate and GWL monitoring data. The analysis of the results has allowed us to study the forecast uncertainty of the applied approach. The results show (see Fig. 4) that it can be used to simulate GWL, providing information about the risk of GWL depletion, and therefore, the necessity of applying pumping rate constraints to define sustainable water resource management in order to avoid undesirable impacts (e.g. in groundwater dependent ecosystems in our case study). Nevertheless, future research should be carried out to study the

Table 1
The best experiments (delay, number of neurons, and ANN approach), number of parameters of the ANN and RMSE, AIC and BIC for the forecasting of each well location.

Piezometer	Delay	Neurons	ANN approach	Number of parameters	RMSE (m)	AIC	BIC
04.04.226	1	1	ELMAN Pef	6	0.73	−52.13	−35.41
04.06.001	2	1	NAR	5	0.23	−310.84	−296.90
04.06.002	9	1	NAR	12	0.70	−59.50	−26.05
04.06.003	1	6	NARX P	25	0.84	30.36	100.04
04.06.004	4	6	NARX Pef	61	1.47	217.72	387.76
04.06.009	3	1	NARX Pef	9	0.45	−167.11	−142.02
04.06.010	10	3	NARX P	67	1.89	288.12	474.88
04.06.011	2	1	ELMAN Pef	7	1.68	140.31	159.82
04.06.012	2	9	NARX Pef	55	0.82	81.51	234.82
04.06.013	1	1	ELMAN P	6	0.44	−182.24	−165.51
04.06.014	2	1	ELMAN Pef	7	0.95	5.72	25.23
04.06.015	7	1	NARX P	17	0.93	23.54	70.92
04.06.016	1	10	NAR	31	0.67	−32.89	53.52
04.06.017	2	1	ELMAN Pef	7	0.82	−32.25	−12.74
04.06.019	1	1	ELMAN Pef	6	0.82	−27.20	−10.47
04.06.020	1	1	ELMAN Pef	6	1.27	84.21	100.94
04.06.021	1	7	NAR	22	0.61	−67.31	−5.99
04.06.023	1	1	NAR	4	2.16	196.84	207.99
04.06.024	1	1	ELMAN Pef	6	1.85	162.02	178.75
04.06.025	9	3	NARX Pef	61	1.05	135.60	305.64
04.06.026	2	3	NARX Pef	19	1.75	174.40	227.36
04.06.027	1	5	NARX Pef	21	0.54	−94.67	−36.13
04.06.028	1	1	ELMAN Pef	6	0.54	−131.61	−114.89
04.06.029	1	1	ELMAN P	6	1.45	104.25	120.97
04.06.030	1	6	NAR	19	0.18	−365.44	−312.47
04.06.031	3	1	NAR	6	0.12	−480.53	−463.80
04.06.032	2	7	NAR	29	0.64	−48.99	31.85
04.06.033	2	1	ELMAN Pef	7	1.00	17.82	37.33
04.06.034	1	1	ELMAN P	6	0.64	−93.10	−76.38
04.06.035	1	5	NARX Pef	21	1.10	65.85	124.38
04.06.112	1	1	NARX P	5	0.15	−439.24	−425.31
04.06.114	10	1	NAR	13	0.19	−362.79	−326.55
04.06.115	2	1	ELMAN Pef	7	0.69	−72.67	−53.16
04.06.201	1	1	ELMAN P	6	1.01	14.98	31.71
04.06.206	1	1	ELMAN Pef	6	1.25	70.91	87.64
04.06.207	1	3	NARX Pef	13	0.54	−118.74	−82.51
04.06.208	2	4	NARX P	25	1.81	205.82	275.51
04.06.209	1	8	NAR	25	0.49	−117.08	−47.39
04.06.211	1	7	NARX Pef	29	0.24	−266.07	−185.23
04.06.213	1	1	ELMAN Pef	6	1.82	160.00	176.73
04.06.214	4	2	NAR	13	0.73	−46.58	−10.34
04.06.216	5	1	NARX P	13	1.03	36.62	72.85
04.06.218	6	1	ELMAN P	11	3.10	298.77	329.44
04.06.221	1	1	ELMAN Pef	6	1.31	78.51	95.24
04.06.224	1	1	ELMAN Pef	6	0.19	−385.25	−368.53
04.06.227	1	1	NAR	4	0.24	−322.81	−311.66
04.06.228	4	1	NARX P	11	0.74	−45.91	−15.25
04.06.229	7	5	NARX Pef	49	0.63	−9.48	127.11
04.06.234	1	1	ELMAN Pef	6	0.34	−236.89	−220.17
04.06.237	5	2	NARX Pef	25	0.63	−48.20	21.49
04.06.238	2	3	NAR	13	0.65	−72.06	−35.83

impact of other sources of uncertainty, in real future horizon forecast, such as the one due to the exogenous variables. These short-term predictions can be even applied to estimate future levels of drought alert in accordance with the historical and forecasted groundwater levels (Pham et al., 2022).

6. Conclusions

In this study, we propose a novel integrated methodological framework for short-term forecasting of groundwater levels by using geostatistics, optimal exogenous variables and artificial neural networks. The methodology is illustrated through the analysis of the aquifer Campo de Montiel (south-eastern Spain). The method is a parsimonious approach (from the point of view of the needed data) that can be applied to any case study.

By using OK for the completion of the GWL data we obtained much better results by using as the target variable the difference between GWL values and the average value than GWL values directly.

In the analysis of optimal exogenous variables, in general Pef showed higher correlations with GWL than P and the accumulation period for which the maximum correlations are reached depends on the well location (it varies from 3 to 30 months). In general, the wells with stronger correlations are located closer to the central valley of the GWB and the wells with weak correlations are located close to the GWB boundaries.

We propose applying the different ANN approaches to each well location in order to find the optimal approach. NAR, which does not consider secondary information from P or Pef, is the best approach for 25.5 % of the cases and it is associated with well locations with lower correlations with P and Pef. The approaches that consider exogenous variables were selected for 74.5 % of the instances. Amongst them, the ones that use Pef have been selected more times as the best experiments. We obtained a mean RMSE of 1.14 m in the test and 0.76, 0.92, 0.92, 0.87, 0.90, and 1.05 m for the forecasting test for the months from 1 to 6 respectively but the accuracy of the results can be

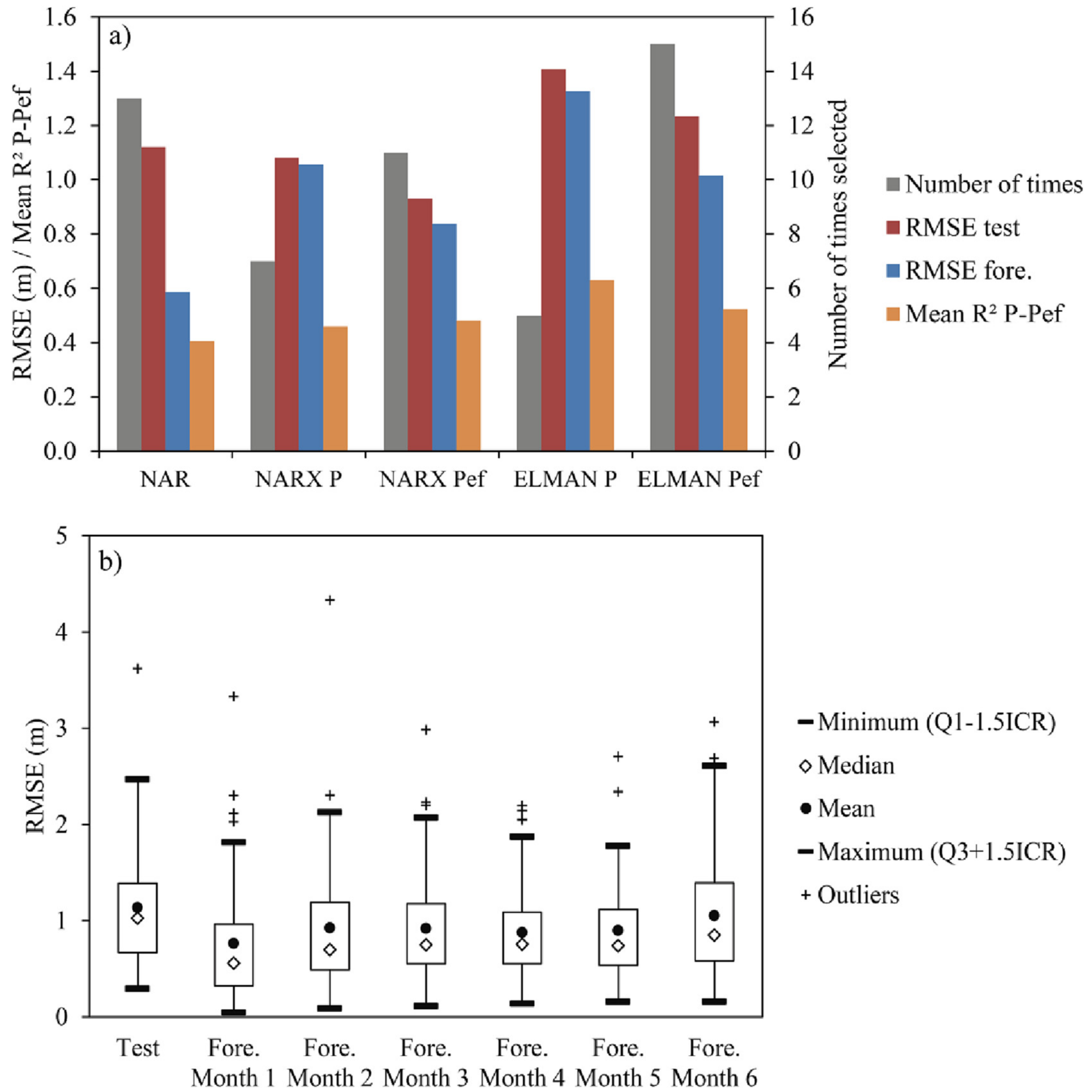


Fig. 12. a) Number of instances when each ANN approach is the best experiment, RMSE of the test, and forecasting test, and mean R^2 with P and Pef and b) box-whiskers of the RMSE for the test and forecasting test for the selected experiments.

considerably different depending on the well location. For each well location and its corresponding best ANN approach and experiment (delay, neurons) we generated multiple estimations (a battery of 200) of the GWL series and forecast in order to take into account the modelling uncertainty.

The proposed methodological framework has shown satisfactory results and is relatively easy to apply, it is operational, and does not have demanding data requirements. Thus, it is recommended for a better understanding of the behaviour of groundwater systems and to forecast the potential future status of the system in order to anticipate management responses.

CRediT authorship contribution statement

A. J. Collados-Lara: Conceptualization, Methodology, Software, Validation, Visualization, Investigation, Data curation, Writing- original draft preparation, Funding acquisition. D. Pulido-Velazquez: Conceptualization, Methodology, Writing- original draft preparation, Funding acquisition. L. G. B. Ruiz: Software, Validation, Writing- reviewing and editing. M. C. Pegalajar: Software, Validation, Writing- reviewing and editing. E. Pardo-Igúzquiza: Methodology, Writing- reviewing and editing. L. Baena-Ruiz: Writing- reviewing and editing.

Funding

This research was partially supported by the Regional Ministry of Economic Transformation, Industry, Knowledge and Universities of the Regional Government of Andalusia through the post-doc programme of the Andalusian Plan for Research Development and Innovation (PAIDI 2021) (POSTDOC_21_00154, University of Granada, Antonio-Juan Collados-Lara), and the research projects SIGLO-AN (RTI2018-101397-B-I00) and SIGLO-PRO (PID2021-128021OB-I00) from the Spanish Ministry of Science and Innovation and the WP4 (Appraisal, protection & sustainable use of Europe's groundwater resources) of the CSA project A Geological Service for Europe (GSEU) (HORIZON-CL5-2021-D3-02-14CSA). Funding for open access charge: Universidad de Granada/CBUA.

Data availability

Data will be made available on request.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

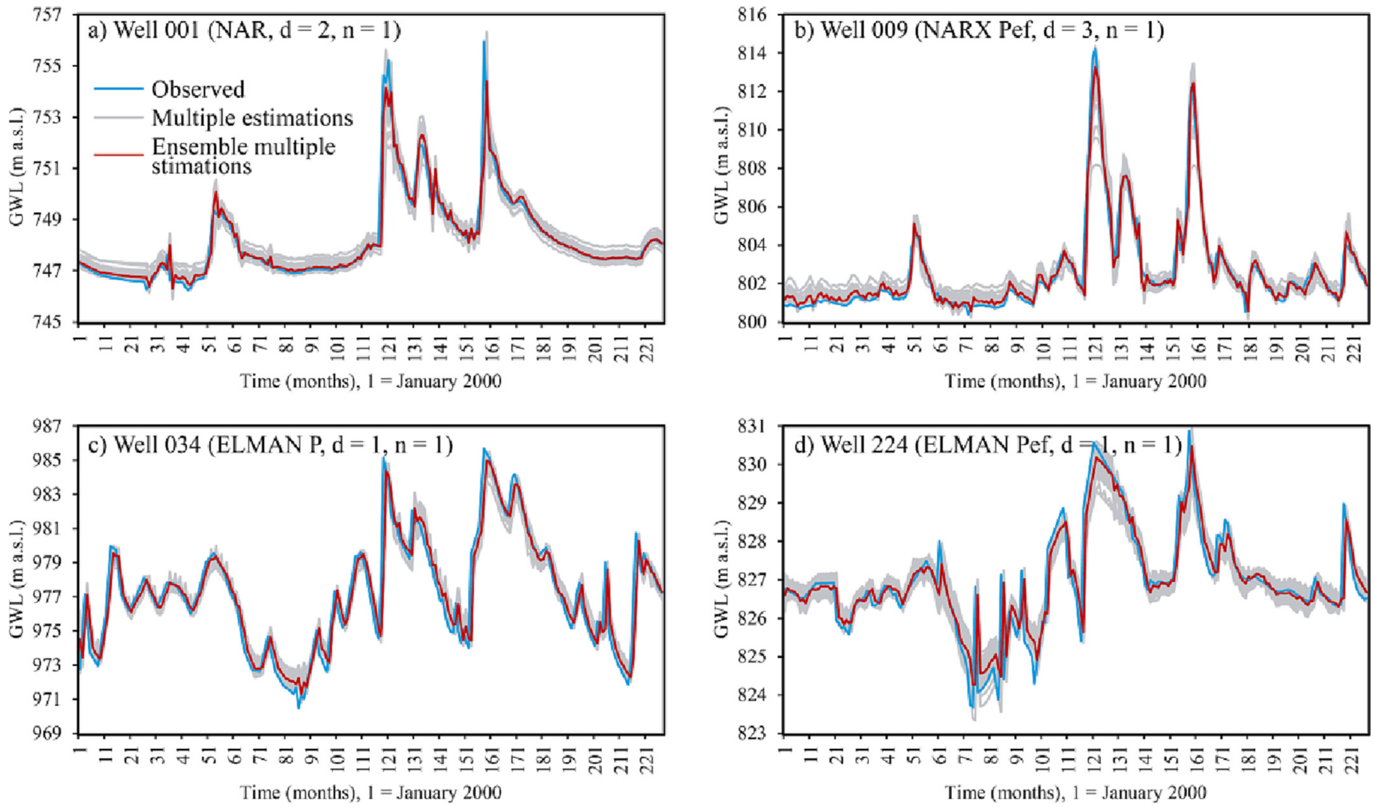


Fig. 13. Observed data of groundwater level, multiple estimated series, and ensemble of the multiple series using the best experiments and ANN approaches for well locations a) 001 (delay = 2, neurons = 1, ANN approach = NAR), b) 009 (delay = 3, neurons = 1, ANN approach = NARX Pef), c) 0034 (delay = 1, neurons = 1, ANN approach = ELMAN P), and d) 224 (delay = 1, neurons = 1, ANN approach = ELMAN Pef).

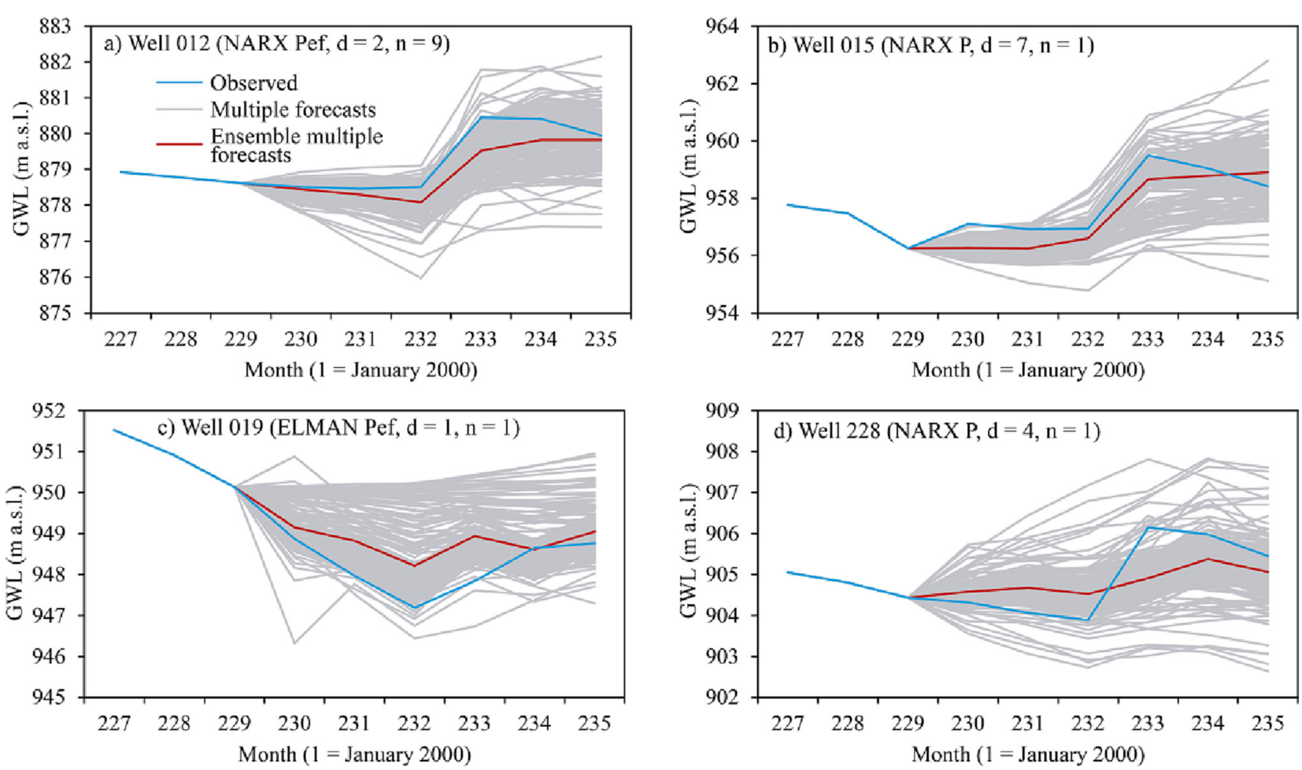


Fig. 14. Observed data of groundwater level, multiple forecast series, and ensemble of the multiple forecast series using the best experiments and ANN approach for well locations a) 012 (delay = 2, neurons = 9, ANN approach = NARX Pef), b) 015 (delay = 7, neurons = 1, ANN approach = NARX P), c) 0019 (delay = 1, neurons = 1, ANN approach = ELMAN Pef), and d) 228 (delay = 4, neurons = 1, ANN approach = NARX P).

References

Adamowski, J., Chan, H.F., 2011. A wavelet neural network conjunction model for groundwater level forecasting. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2011.06.013.

Aguilera, H., Guardiola-Albert, C., Naranjo-Fernández, N., Kohfahl, C., 2019. Towards flexible groundwater-level prediction for adaptive water management: using Facebook's prophet forecasting approach. Hydrol. Sci. J. https://doi.org/10.1080/02626667.2019.1651933.

Ahmadi, A., Emami, M., Daccache, A., He, L., 2021. Soil properties prediction for precision agriculture using visible and near-infrared spectroscopy: a systematic review and meta-analysis. Agronomy https://doi.org/10.3390/agronomy11030433.

Baena-Ruiz, L., Pulido-Velazquez, D., Collados-Lara, A.J., Renau-Pruñonosa, A., Morell, I., 2018. Global assessment of seawater intrusion problems (Status and Vulnerability). Water Resour. Manag. https://doi.org/10.1007/s11269-018-1952-2.

Bhat, S., Motz, L.H., Pathak, C., Kuebler, L., 2015. Geostatistics-based groundwater-level monitoring network design and its application to the Upper Floridan aquifer, USA. Environ. Monit. Assess. https://doi.org/10.1007/s10661-014-4183-x.

Barker, L.J., Hannaford, J., Chiverton, A., Svensson, C., 2016. From meteorological to hydrological drought using standardised indicators. Hydrol. Earth Syst. Sci. https://doi.org/10.5194/hess-20-2483-2016.

Chen, C., He, W., Zhou, H., Xue, Y., Zhu, M., 2020. A comparative study among machine learning and numerical models for simulating groundwater dynamics in the Heihe River basin, northwestern China. Sci. Rep. 10 (1). https://doi.org/10.1038/s41598-020-60698-9.

Chilès, J.P., Delfiner, P., 1999. Geostatistics, Modeling Spatial Uncertainty. Wiley.

Collados-Lara, A.J., Gomez-Gomez, J.D.D., Pulido-Velazquez, D., Pardo-Iguzquiza, E., 2022. An approach to identify the best climate models for the assessment of climate change impacts on meteorological and hydrological droughts. Nat. Hazards Earth Syst. Sci. https://doi.org/10.5194/nhess-22-599-2022.

Collados-Lara, A.J., Pardo-Igúzquiza, E., Pulido-Velazquez, D., Baena-Ruiz, L., 2021. Estimation of the monthly dynamics of surface water in wetlands from satellite and secondary hydro-climatological data. Remote Sens. https://doi.org/10.3390/rs13122380.

Collados-Lara, A.J., Pardo-Igúzquiza, E., Pulido-Velazquez, D., Jiménez-Sánchez, J., 2018. Precipitation fields in an alpine Mediterranean catchment: inversion of precipitation gradient with elevation or undercatch of snowfall? Int. J. Climatol. https://doi.org/10.1002/joc.5517.

Collados-Lara, A.J., Pulido-Velazquez, D., Pardo-Iguzquiza, E., 2020. A statistical tool to generate potential future climate scenarios for hydrology applications. Sci. Program. https://doi.org/10.1155/2020/8847571.

Coppola, E.A., Rana, A.J., Poulton, M.M., Szidarovszky, F., Uhl, V.W., 2005. A neural network model for predicting aquifer water level elevations. Ground Water https://doi.org/10.1111/j.1745-6584.2005.0003.x.

Cramer, W., Guiot, J., Fader, M., Garrabou, J., Gattuso, J.P., Iglesias, A., Xoplaki, E., 2018. Climate change and interconnected risks to sustainable development in the Mediterranean. Nat. Clim. Chang. https://doi.org/10.1038/s41558-018-0299-2.

Dashti, Z., Nakhaei, M., Vadiati, M., Karami, G.H., Kisi, O., 2023. A literature review on pumping test analysis (2000–2022). Environ. Sci. Pollut. Res. 30 (4), 9184–9206. https://doi.org/10.1007/s11356-022-24440-4.

Di Nunno, F., Granata, F., 2020. Groundwater level prediction in Apulia region (Southern Italy) using NARX neural network. Environ. Res. https://doi.org/10.1016/j.envres.2020.110062.

Didovets, I., Krysanova, V., Bürger, G., Snizhko, S., Balabukh, V., Bronstert, A., 2019. Climate change impact on regional floods in the carpathian region. Journal of hydrology. Reg. Stud. https://doi.org/10.1016/j.ejrh.2019.01.002.

Dubois, E., Larocque, M., Gagné, S., Braun, M., 2022. Climate change impacts on groundwater recharge in cold and humid climates: controlling processes and thresholds. Climate https://doi.org/10.3390/cli10010006.

Essenfelder, A.H., Giupponi, C., 2020. A coupled hydrologic-machine learning modelling framework to support hydrologic modelling in river basins under interbasin water transfer regimes. Environ. Model. Softw. 131. https://doi.org/10.1016/j.envsoft.2020.104779.

Gomez-Gomez, J.de D., Pulido-Velazquez, D., Collados-Lara, A.J., Fernandez-Chacon, F., 2022. The impact of climate change scenarios on droughts and their propagation in an arid Mediterranean basin. A useful approach for planning adaptation strategies. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2022.153128.

Goodfellow, I., Bengio, Y., Courville, A., 2016. Deep learning. MIT press. https://www.deeplearningbook.org/.

Guzman, S.M., Paz, J.O., Tagert, M.L.M., 2017. The use of NARX neural networks to forecast daily groundwater levels. Water Resour. Manag. https://doi.org/10.1007/s11269-017-1598-5.

Hargreaves, G.H., Allen, R.G., 2003. History and evaluation of hargreaves evapotranspiration equation. J. Irrig. Drain. Eng. https://doi.org/10.1061/(asce)0733-9437(2003)129:1(53).

Herath, H.M.V.V., Chadalawada, J., Babovic, V., 2021. Hydrologically informed machine learning for rainfall-runoff modelling: towards distributed modelling. Hydrol. Earth Syst. Sci. 25 (8). https://doi.org/10.5194/hess-25-4373-2021.

Herrera, S., Fernández, J., Gutiérrez, J.M., 2016. Update of the Spain02 gridded observational dataset for EURO-CORDEX evaluation: assessing the effect of the interpolation methodology. Int. J. Climatol. https://doi.org/10.1002/joc.4391.

Hidalgo-Hidalgo, J.-D., Collados-Lara, A.-J., Pulido-Velazquez, D., Rueda, F.J., Pardo-Igúzquiza, E., 2022. Analysis of the potential impact of climate change on climatic droughts, snow dynamics, and the correlation between them. Water https://doi.org/10.3390/w14071081.

Hussein, E.A., Thron, C., Ghaziasgar, M., Bagula, A., Vaccari, M., 2020. Groundwater prediction using machine-learning tools. Algorithms https://doi.org/10.3390/a13110300.

Jimeno-Sáez, P., Senent-Aparicio, J., Pérez-Sánchez, J., Pulido-Velazquez, D., María Cecilia, J., 2017. Estimation of instantaneous peak flow using machine-learning models and empirical formula in peninsular Spain. Water (Switzerland) https://doi.org/10.3390/w9050347.

Kay, A.L., Rudd, A.C., Fry, M., Nash, G., Allen, S., 2021. Climate change impacts on peak river flows: combining national-scale hydrological modelling and probabilistic projections. Clim. Risk Manag. 31. https://doi.org/10.1016/j.crm.2020.100263.

Kotlarski, S., Szabó, P., Herrera, S., Räty, O., Keuler, K., Soares, P.M., Pianko-Kluczyńska, K., 2019. Observational uncertainty and regional climate model evaluation: a pan-european perspective. Int. J. Climatol. https://doi.org/10.1002/joc.5249.

Leduc, C., Pulido-Bosch, A., Remini, B., 2017. Anthropization of groundwater resources in the Mediterranean region: processes and challenges. Hydrogeol. J. https://doi.org/10.1007/s10040-017-1572-6.

Linés, C., Iglesias, A., Garrote, L., Sotés, V., Werner, M., 2018. Do users benefit from additional information in support of operational drought management decisions in the Ebro basin? Hydrol. Earth Syst. Sci. https://doi.org/10.5194/hess-22-5901-2018.

Llopis-Albert, C., Pulido-Velazquez, D., 2015. Using MODFLOW code to approach transient hydraulic head with a sharp-interface solution. Hydrol. Process. https://doi.org/10.1002/hyp.10354.

Martinsen, G., Bessiere, H., Caballero, Y., Koch, J., Collados-Lara, A.J., Mansour, M., Stisen, S., 2022. Developing a pan-european high-resolution groundwater recharge map – combining satellite data and national survey data using machine learning. Sci. Total Environ. https://doi.org/10.1016/j.scitotenv.2022.153464.

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. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2015.10.018.

Mengistu, T.D., Chung, I.M., Chang, S.W., Yifru, B.A., Kim, M.G., Lee, J., Kim, I.H., 2021. Challenges and prospects of advancing groundwater research in ethiopian aquifers: a review. Sustainability (Switzerland) https://doi.org/10.3390/su132011500.

Moseki, M.C., 2018. Climate change impacts on groundwater: literature review. Environ. Risk Assess. Remediat. https://doi.org/10.4066/2529-8046.100033.

Mukherjee, S., Mishra, A., Trenberth, K.E., 2018. Climate change and drought: a perspective on drought indices. Curr. Clim. Chang. Rep. https://doi.org/10.1007/s40641-018-0098-x.

Nourani, V., Ghaneei, P., Kantoush, S.A., 2022. Robust clustering for assessing the spatiotemporal variability of groundwater quantity and quality. J. Hydrol. 604. https://doi.org/10.1016/j.jhydrol.2021.127272.

Nourani, V., Mousavi, S., 2016. Spatiotemporal groundwater level modeling using hybrid artificial intelligence-meshless method. J. Hydrol. 536. https://doi.org/10.1016/j.jhydrol.2016.02.030.

Pardo-Igúzquiza, E., Collados-Lara, A.J., Pulido-Velazquez, D., 2019. Potential future impact of climate change on recharge in the sierra de las nieves (southern Spain) high-relief karst aquifer using regional climate models and statistical corrections. Environ. Earth Sci. https://doi.org/10.1007/s12665-019-8594-4.

Pham, Q.B., Kumar, M., Di Nunno, F., Elbeltagi, A., Granata, F., Islam, A.R.M.T., Anh, D.T., 2022. Groundwater level prediction using machine learning algorithms in a drought-prone area. Neural Comput. & Applic. https://doi.org/10.1007/s00521-022-07009-7.

Pulido-Velazquez, D., Ahlfeld, D., Andreu, J., Sahuquillo, A., 2008. Reducing the computational cost of unconfined groundwater flow in conjunctive-use models at basin scale assuming linear behaviour: the case of Adra-campo de Dalías. J. Hydrol. https://doi.org/10.1016/j.jhydrol.2008.02.006.

Pulido-Velazquez, D., Baena-Ruiz, L., Mayor, B., Zorrilla-Miras, P., López-Gunn, E., de Dios Gómez-Gómez, J., de la Hera-Portillo, Á., Collados-Lara, A.-J., Moreno, M.M., Aróstegui, J.L.G., Alcalá, F.J., 2023. Integrating stakeholders' inputs to co-design climate resilience adaptation measures in Mediterranean areas with conflicts between wetland conservation and intensive agriculture. Sci. Total Environ. 870, 161905. https://doi.org/10.1016/j.scitotenv.2023.161905.

Pulido-Velazquez, D., Romero, J., Collados-Lara, A.J., Alcalá, F.J., Fernández-Chacón, F., Baena-Ruiz, L., 2020. Using the turnover time index to identify potential strategic groundwater resources to manage droughts within continental Spain. Water (Switzerland) https://doi.org/10.3390/w12113281.

Pulido-Velazquez, D., Renau-Pruñonosa, A., Llopis-Albert, C., Morell, I., Collados-Lara, A.J., Senent-Aparicio, J., Baena-Ruiz, L., 2018. Integrated assessment of future potential global change scenarios and their hydrological impacts in coastal aquifers - a new tool to analyse management alternatives in the plana Oropesa-Torreblanca aquifer. Hydrol. Earth Syst. Sci. https://doi.org/10.5194/hess-22-3053-2018.

Razavi, S., Tolson, B.A., Burn, D.H., 2012. Review of surrogate modeling in water resources. Water Resour. Res. 48 (7). https://doi.org/10.1029/2011WR011527.

Rohde, M.M., Biswas, T., Housman, I.W., Campbell, L.S., Klausmeyer, K.R., Howard, J.K., 2021. A machine learning approach to predict groundwater levels in California reveals ecosystems at risk. Front. Earth Sci. 9. https://doi.org/10.3389/feart.2021.784499.

Roshni, T., Jha, M.K., Drisya, J., 2020. Neural network modeling for groundwater-level forecasting in coastal aquifers. Neural Comput. & Applic. https://doi.org/10.1007/s00521-020-04722-z.

Senent-Aparicio, J., López-Ballesteros, A., Pérez-Sánchez, J., Segura-Méndez, F.J., Pulido-Velazquez, D., 2018. Using multiple monthly water balance models to evaluate gridded precipitation products over peninsular Spain. Remote Sens. https://doi.org/10.3390/rs10060922.

Shen, C., Lawson, K., 2021. Applications of deep learning in hydrology. Deep Learning for the Earth Sciences https://doi.org/10.1002/9781119646181.ch19.

Shen, C., Chen, X., Laloy, E., 2021. Editorial: broadening the use of machine learning in hydrology. Frontiers in Water. Vol. 3. https://doi.org/10.3389/frwa.2021.681023.

Tao, H., Hameed, M.M., Marhoon, H.A., Zounemat-Kermani, M., Heddam, S., Sungwon, K., Yaseen, Z.M., 2022. Groundwater level prediction using machine learning models: a comprehensive review. Neurocomputing https://doi.org/10.1016/j.neucom.2022.03.014.

Tramblay, Y., Koutroulis, A., Samaniego, L., Vicente-Serrano, S.M., Volaire, F., Boone, A., Polcher, J., 2020. Challenges for drought assessment in the Mediterranean region under future climate scenarios. Earth Sci. Rev. https://doi.org/10.1016/j.earscirev.2020.103348.

Wunsch, A., Liesch, T., Broda, S., 2021. Groundwater level forecasting with artificial neural networks: a comparison of long short-term memory (LSTM), convolutional neural networks (CNNs), and non-linear autoregressive networks with exogenous input (NARX). Hydrol. Earth Syst. Sci. https://doi.org/10.5194/hess-25-1671-2021.

Yang, Y., Chui, T.F.M., 2021. Reliability assessment of machine learning models in hydrological predictions through metamorphic testing. Water Resour. Res. 57 (9). https://doi.org/10.1029/2020WR029471.

Xu, T., Valocchi, A.J., Choi, J., Amir, E., 2014. Use of machine learning methods to reduce predictive error of groundwater models. Groundwater https://doi.org/10.1111/gwat.12061.

Zaadnoordijk, W.J., Bus, S.A.R., Lourens, A., Berendrecht, W.L., 2019. Automated time series modeling for piezometers in the National Database of the Netherlands. Groundwater https://doi.org/10.1111/gwat.12819.

Zhang, R., Zen, R., Xing, J., Arsa, D.M.S., Saha, A., Bressan, S., 2020. Hydrological Process Surrogate Modelling and Simulation with Neural Networks. Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics), 12085 LNAI. https://doi.org/10.1007/978-3-030-47436-2_34.