Multistep Ahead Groundwater Level Time-Series Forecasting Using Gaussian Process Regression and ANFIS

N. Sujay Raghavendra and Paresh Chandra Deka

Abstract Groundwater level is regarded as an environmental indicator to quantify groundwater resources and their exploitation. In general, groundwater systems are characterized by complex and nonlinear features. Gaussian Process Regression (GPR) approach is employed in the present study to investigate its applicability in probabilistic forecasting of monthly groundwater level fluctuations at two shallow unconfined aquifers located in the Kumaradhara river basin near Sullia Taluk, India. A series of monthly groundwater level observations monitored during the period 2000–2013 is utilized for the simulation. Univariate time-series GPR and Adaptive Neuro Fuzzy Inference System (ANFIS) models are simulated and applied for multistep lead time forecasting of groundwater levels. Individual performance of the GPR and ANFIS models are comparatively evaluated using various statistical indices. In overall, simulation results reveal that GPR model provided reasonably accurate predictions than that of ANFIS during both training and testing phases. Thus, an effective GPR model is found to generate more precise probabilistic forecasts of groundwater levels.

Keywords ANFIS · Groundwater system · Gaussian process regression · Time-series forecasting

1 Introduction

Over the past decade, groundwater depletion is one of the major issues worldwide, which is posing direct or indirect impacts on human livelihoods, flora and fauna, natural habitat, and ecosystems. Depletion of groundwater storage, land subsidence,

N.S. Raghavendra (✉) · P.C. Deka
Department of Applied Mechanics and Hydraulics, National Institute
of Technology Karnataka, Surathkal, Mangalore 575025, India
e-mail: sujayraghavendran@ymail.com

P.C. Deka
e-mail: pareshdeka@yahoo.com

reductions in stream flow and lake water levels, saltwater intrusion, loss of wetland and riparian ecosystems, and variations in groundwater quality are some of the vital factors influencing the sustainability of groundwater resources [1]. Sustainable groundwater resources development is the key issue to be addressed by policy makers or water managers by implementing various alternative management strategies. Groundwater restoration or recycle is not equally fast as that of surface water; it may take place after many years. Thus, constant monitoring of groundwater levels is extremely important for reliable assessment of temporal availability of groundwater at any required location [2]. The benefits of groundwater level forecasting include assessment of annual and long-term changes in groundwater storage, estimation of recharge rates, manage drinking water demand, and to ensure the sustainable use of groundwater resources [3].

Till date, several deterministic, stochastic and time-series based models have been developed for the forecasting of groundwater levels [4–7]. In the recent past, soft computing tools like Artificial Neural Network (ANN), ANFIS, Support Vector Regression (SVR), and so on have also been widely utilized for groundwater level prediction studies [8–14]. Quite a few hybrid artificial intelligence models developed by incorporating wavelet analysis efficiently forecast groundwater levels at different time scales [15–18]. Determining a model which is capable to efficiently capture the nonlinearities of the data without overfitting is the crucial job while modeling using time-series data. The ability to select the hyper parameters of the kernel automatically is one of the prominent benefits of Gaussian processes over conventional kernel interpretations of regression. The Bayesian learning algorithm-based Gaussian Process Regression is successfully applied for prediction of nonstationary time-series [19], monthly stream flow forecasting [20], and stream water temperature prediction [21], and so on. Compared to the conventional time-series forecasting methods, GPR model is said to possess strong nonlinear mapping ability, estimation of uncertainty, and is greatly fault-tolerant [22, 23]. Hence, in this paper, we demonstrate the state-of-the-art capability of Gaussian process regression for multistep lead time probabilistic forecast of groundwater level fluctuations. The ANFIS model is also employed for comparative study with GPR forecasts.

2   Study Area and Data Analysis

The study area (Fig. 1) is located near to southwest coast in the state of Karnataka, India. The observation wells selected for the current study are located inside the Kumaradhara river basin which covers a geographical area of 1776 sq km and is located in between 12º 29′ 04″ and 12º 58′ 33″ north latitude and from 75º 09′ 58″ to 75º 47′ 48″ east longitudes. The observation well located at Bellare lies at 12° 39′ 53″ north latitude and 75° 17′ 18″ east longitude, while the other well at Guttigaru lies at 12° 37′ 53″ north latitude and 75° 31′ 44″ east longitude as shown in Fig. 1.
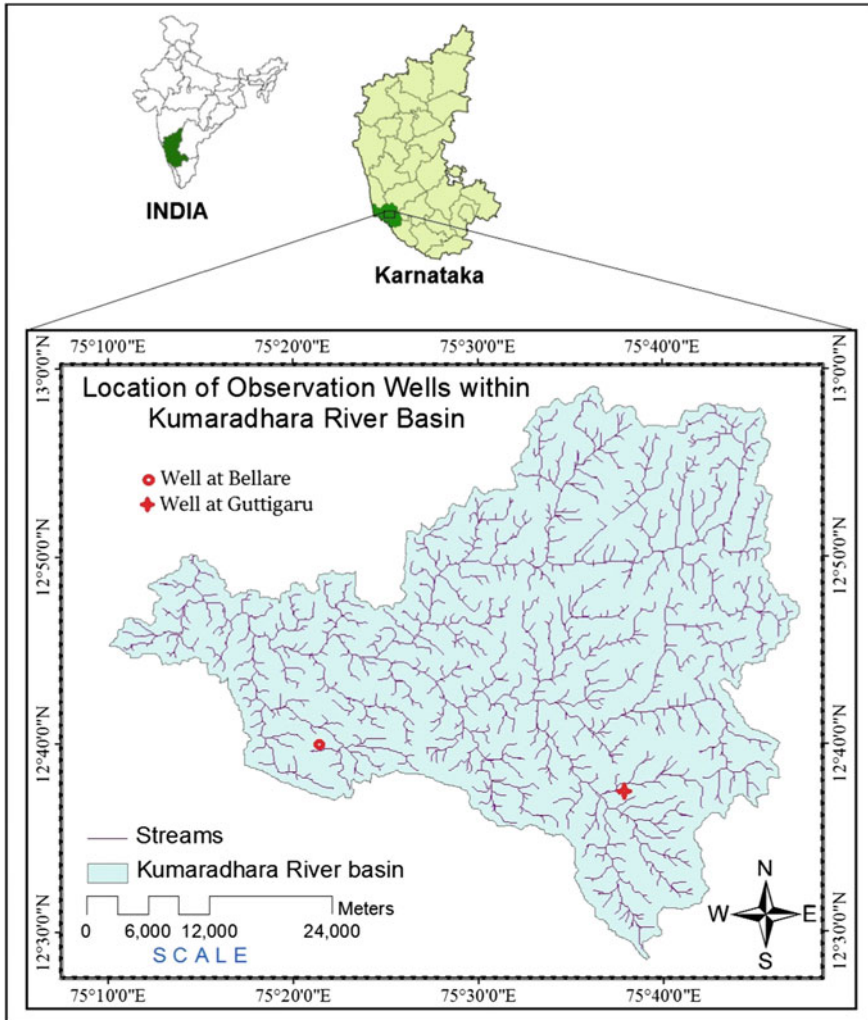
Fig. 1 Study area (location of observation wells)

The study area has a tropical monsoon climate dominated by the southwest monsoon (June–October). The mean annual rainfall over the basin is around 3,500 mm. The geology of the area is predominantly characterized by Lateritic soil with highly porous and permeable nature. Due to this lateritic soil property, shallow groundwater levels in the selected unconfined aquifers follow a regular cyclic pattern of seasonal fluctuation, typically rising during the monsoon due to greater precipitation and recharge, then declining during the summer.

The groundwater level data of the observation wells located at Bellare and Guttigaru for the years 2000–2013 were retrieved from Department of Mines

Table 1 Statistical parameters of groundwater level dataset

Groundwater table below ground level (m)				
Well location	Bellare		Guttigaru	
Dataset	Train	Test	Train	Test
X_{max}	13.05	12.13	11.24	12.38
X_{min}	4.28	5.21	1.33	3.2
X_{mean}	9.25	8.53	7.11	7.37
X_{mode}	10.18	5.58	4.27	3.29
S_d	2.25	2.26	2.98	3.15
C_v	0.24	0.27	0.42	0.43

and Geology, Dakshina Kannada Dist., Govt. of Karnataka, India. The topographic elevation of these wells is about 100–130 m above mean sea level. This data comprises of 166 monthly observations, in which 10 years of data during Jan 2000–Dec 2009 is used for model training and remaining 4 years of data during Jan 2010–Oct 2013 is used as out-of-sample set or testing set to measure the predictability of the developed models.

The descriptive statistics of the observed groundwater levels in the two observation wells are presented in Table 1. The X_{max}, X_{min}, X_{mean}, X_{mode}, S_d, and C_v denotes the maximum, minimum, mean, mode, standard deviation and coefficient of variation respectively. Since the mean and coefficient of variation of the observed groundwater level dataset don't vary ominously during training and testing periods, it could be inferred as a reasonable stationary time-series. In the present scenario, GPR and ANFIS models are explored to forecast 1, 3 and 6 months ahead groundwater level fluctuations. Monthly groundwater level time-series up to previous four time steps are taken as input variables. In order to test the hypothesis that $GWL_{(t-2)},...GWL_{(t-p)}$ further help in forecasting $GWL_{(t)}$, beyond $GWL_{(t-1)}$, one can use an F-test. The lag order $p = 4$ was determined from the F-test statistic. F-test is the test statistic to examine the significance of the components in the model [24]. The expected output from the developed models is the groundwater level at time step t, $t + 3$, and $t + 6$. The input-output combinations are as presented below.

I. $GWL_{(t-4)} + GWL_{(t-3)} + GWL_{(t-2)} + GWL_{(t-1)} = GWL_{(t)}$
II. $GWL_{(t-4)} + GWL_{(t-3)} + GWL_{(t-2)} + GWL_{(t-1)} = GWL_{(t+3)}$
III. $GWL_{(t-4)} + GWL_{(t-3)} + GWL_{(t-2)} + GWL_{(t-1)} = GWL_{(t+6)}$

3 Methodology

In the present study, Gaussian Process Regression (GPR) and Adaptive Neuro Fuzzy Inference System (ANFIS) approaches are proposed for model development of groundwater level time-series forecasting. GPR and ANFIS is used for 1, 3, and 6 month lead groundwater level time-series forecasting using lagged input data up to 4 months in the past.

3.1 Gaussian Process Regression

Gaussian process regression is a standard method in probability theory wherein the interpolated values are modeled by a Gaussian process governed by prior covariance. Incorporating appropriate assumptions on the priors, GPR renders the best linear unbiased prediction of the values [25]. GPs constitute one of the most important Bayesian discriminative kernel learning approach due to its practical and theoretical simplicity and outstanding generalization ability. A sequence of random variables $\{X_n\}$ defining a stationary process can have any probability distribution. A stationary process $\{X_n\}$ is called a Gaussian process, if the joint distribution of $(X_{n+1}, X_{n+2},\ldots, X_{n+k})$ is a k-variate normal for every positive integer k.

Consider an observation space χ. A GP $f(x)$, where $x \in \chi$, is defined by a set of random variables, any finite number of which possess a joint Gaussian distribution function which is fully specified by its mean function $m(x)$ and covariance $k(x, x')$ [26].

So let,

$$
\begin{aligned}
m(x) &= E[f(x)] \\
k(x,x') &= E[(f(x) - m(x)) \cdot (f(x') - m(x))]
\end{aligned}
\tag{1}
$$

Now we can write GP as

$$
f(x) \sim \mathbb{N}(m(x), k(x,x'))
\tag{2}
$$

Consider a training set $D = \{(x_i, y_i) | i = 1, 2, \ldots, N\}$, with m-dimensional input variables, x_i being the observed data related to the phenomenon that is to be modeled and scalars y_i being the associated target values given by $y_i = f(x_i) + \epsilon_i$, where ϵ_i is Gaussian noise with zero mean and variance σ_n^2.

The joint normality of the training target values $y = [y_i]_{i=1}^N$ and some unknown target value y_*, are estimated by the value f_* of the hypothesized GP assessed at the observation point x_*, yields

$$
\begin{bmatrix} y \\ f_* \end{bmatrix} \sim \mathbb{N}\left(0, \begin{bmatrix} K(X,X) + \sigma_N^2 I_N & k(x_*) \\ k(x_*)^{\mathrm{T}} & k(x_*, x_*) \end{bmatrix}\right)
\tag{3}
$$

where,

$$
k(x_*) \triangleq [k(x_1, x_*), \ldots, k(x_N, x_*)]^T
\tag{4}
$$

$X = [x_i]_{i=1}^N$, I_N, is the $N \times N$ identity matrix, $k(x_*)$ is the vector of covariance between f_* and the training latent function values, and K is the matrix of the covariance between the N training data points (design matrix)

$$K[X,X] \triangleq \begin{bmatrix} k(x_1,x_1) & k(x_1,x_2) & \dots\dots & k(x_1,x_N) \\ k(x_2,x_1) & k(x_2,x_2) & \dots\dots & k(x_2,x_N) \\ . & . & & . \\ . & . & & . \\ k(x_N,x_1) & k(x_N,x_2) & \dots\dots & k(x_N,x_N) \end{bmatrix} \qquad (5)$$

Then, from (Eq. 3) and conditioning on the available training samples, we can derive the expression for the model predictive distribution, yielding

$$p(f_*|x_*,D) = \mathbb{N}(f_*|\mu_*,\sigma_*^2), \qquad (6)$$

where

$$\begin{cases} \mu_* = k(x_*)^{\mathrm{T}}\big(K(X,X) + \sigma_N^2 I_N\big)^{-1} \times y \\ \sigma_*^2 = \sigma_N^2 - k(x_*)^{\mathrm{T}}\big(K(X,X) + \sigma_N^2 I_N\big)^{-1} k(x_*) + k(x_*,x_*) \end{cases} \qquad (7)$$

The covariance function is parameterized by optimal value of hyper parameters. The predictive variance of the GP model is as given in Eq. (7), and it does not depend on the training target values, but depends only on the training input values [27]. The optimal value of hyper parameters of a Gaussian process with any kernel θ, for any distinct data set can be derived by maximizing the log marginal likelihood by means of general optimization procedures. The log marginal likelihood function under the GPR model is presented in Eq. 8 given below.

$$\left\{ \begin{aligned} \log p\big(y|X;\theta,\sigma^2\big) &= -\frac{N}{2}\log 2\pi - \frac{1}{2}\log\big|K(X,X) + \sigma_N^2 I_N\big| \\ &\quad - \frac{1}{2}y^{\mathrm{T}}\big(K(X,X) + \sigma_N^2 I_N\big)^{-1} y \end{aligned} \right\} \qquad (8)$$

3.2 Adaptive Neuro Fuzzy Inference System (ANFIS)

ANFIS is the fuzzy-logic based paradigm integrated with the learning power of Artificial Neural Network (ANN) to improve the intelligent system's performance utilizing knowledge acquired after learning. For a given input-−output data set, ANFIS constructs a hybrid learning algorithm that associates the backpropagation gradient descent and least squares methods to frame a fuzzy inference system whose membership function (MF) parameters are iteratively tuned or adjusted. Adaptive Neuro Fuzzy inference systems comprise of mainly five layers–rule base, database, fuzzification interface, defuzzification interface and decision-making unit. The generalized ANFIS architecture proposed by Jang [28] is summarized below (Fig. 2).
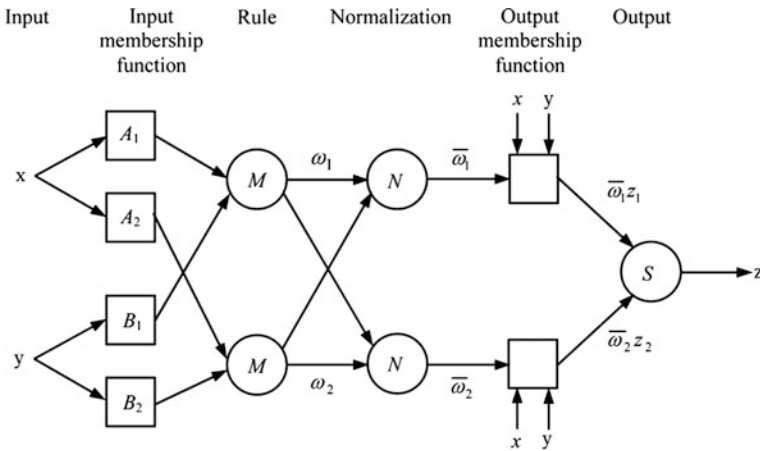
Fig. 2 General ANFIS architecture with two membership functions on each of the two inputs

ANFIS architecture comprises of five layers. Every single node in layer 1 is an adaptive node with a node function which may be anyone among the membership functions. Every node of layer 2 is a fixed node labeled '*M*' which signposts the firing strength of each rule. All nodes of layer 3 are fixed nodes labeled as '*N*' which demonstrates the normalized firing strength of each rule. The Layer 4 is as similar to layer 1 wherein every node is an adaptive node governed by a node function. The layer 5 being a single fixed node labeled '*S*', representing the overall output (z), defined as the summation of all incoming signals [29].

In the present study, we examine three types of membership functions (MFs) namely trapezoidal, gaussian, and generalized bell. Among all the three types of the MFs, we impart two MFs on each of our four inputs, in which eight altogether. With this, the FIS structure consists of 16 fuzzy rules with 104 parameters. A hybrid algorithm integrating the least squares method and the backpropagation gradient descent method is applied to optimize and adjust the generalized bell membership function parameters and coefficients of the output linear equations. The number of epochs and error tolerance is set to 1000 and 0, respectively. From the result as presented in Table 3, it is determined that the ANFIS structure with Generalized bell MF to be better performing than Trapezoidal and Gaussian shaped MFs based on the performance evaluation using correlation coefficient statistic as mentioned below in Sect. 4. Hence, generalized bell MF-based ANFIS models are developed for all the 1, 3 and 6 month lead time forecasting scenarios.

4   Performance Evaluation

The following statistical indices are used to evaluate the performance of both the GPR and ANFIS models in forecasting groundwater level time-series.

$$CC = \frac{\sum\limits_{i=1}^{N}\left\{\left(X_i - \overline{X}\right) \cdot \left(Y_i - \overline{Y}\right)\right\}}{\sqrt{\sum\limits_{i-1}^{N}\left\{\left(X_i - \overline{X}\right)^2 \cdot \left(Y_i - \overline{Y}\right)^2\right\}}} \qquad (9)$$

$$RMSE = \sqrt{\frac{\sum\limits_{i=1}^{N}(X_i - Y_i)}{N}} \qquad (10)$$

$$NSE = 1 - \frac{\sum\limits_{i=1}^{N}(X_i - Y_i)^2}{\sum\limits_{i=1}^{N}\left(X_i - \overline{X}\right)^2} \qquad (11)$$

where,
CC   Correlation Coefficient;
RMSE   Root Mean Squared Error;
NSE   Nash-Sutcliffe Efficiency;
X   Observed/Actual values;
Y   Modeled/Computed values;
$\overline{X}$ Mean of Actual data.

5   Results and Discussion

An appealing characteristic of time-series modeling is that it is based on relatively few assumptions which usually lead to yield good fits. The GPR package in the WEKA 3.6 software [30] is employed to develop the GPR models. The GPR employing Pearson VII function-based universal kernel (PuK) is used for model development. The GPR model developed in the present study is propelled to provide better groundwater level forecasting results. Table 2 presents the developed GP regression model equations. The statistical adequacies of the GPR and ANFIS models for 1, 3 and 6 month ahead forecasts are summarized in Tables 4, 5, and 6, respectively. For both study sites (Bellare and Guttigaru), the GPR models are found to provide more accurate groundwater level forecasts than that of ANFIS

Table 2 Values of Gaussian process regression equations

GPR forecast	Average target value	Inverted covariance matrix		Inverted covariance matrix × target-value vector	
		Lowest value	Highest value	Lowest value	Highest value
Groundwater monitoring well near Bellare					
1 Month lead	9.1686	−0.2107	0.9420	−4.2888	2.1221
3 Month lead	9.1978	−0.2107	0.9420	−4.6319	3.5733
6 Month lead	9.2933	−0.2107	0.9420	−5.0088	3.6627
Groundwater monitoring well near Guttigaru					
1 Month lead	7.0066	−0.1949	0.9179	−4.4780	3.4013
3 Month lead	7.0806	−0.1949	0.9179	−3.6057	5.1893
6 Month lead	7.2328	−0.1949	0.9179	−4.3094	4.5839

Table 3 Performance of ANFIS models at 1 month lead time forecasting

Correlation coefficient (CC)				
ANFIS model with	Bellare		Guttigaru	
	Train	Test	Train	Test
Trapezoidal MF	0.6	0.54	0.58	0.55
Gaussian MF	0.81	0.74	0.79	0.71
Generalized Bell MF	0.95	0.85	0.89	0.83

models for 1, 3 and 6 month lead time forecasting. The GPR models for the Bellare and Guttigaru well sites have a testing RMSE of 0.632 and 1.05 m, respectively (Table 4), and are superior to the ANFIS model forecast, which has a testing RMSE of 0.742 m for the Bellare well site and 1.39 m for the Guttigaru well site during 1 month lead time forecasting (Table 3).

It can be observed (from Tables 4, 5 and 6) that the correlation coefficients of both the GPR and ANFIS models are high during training (calibration). However, during the testing phase, the GPR model is better when compared to ANFIS model. It is noteworthy that the GPR model shows enhanced performance in contrast to ANFIS model, in case of both the wells. The RMSE statistic of multistep lead time forecasting is presented in Fig. 3 wherein it can be inferred that the GPR and ANFIS models are more capable in the shorter lead time forecast. It can be seen that the forecasting efficiency declines during longer lead time forecast. The ANFIS model performs marginally similar to GPR model for 1 month ahead groundwater level forecasting, but for the higher lead times, such as 3 and 6 month lead time, GPR performance is observed better than ANFIS model results as presented in Tables 4, 5 and 6.

Table 4 Performance of GPR and ANFIS models during 1 month lead time forecasting

Well location		Bellare			Guttigaru		
Statistical indices		RMSE (m)	NSE	CC	RMSE (m)	NSE	CC
GPR	TRAIN	0.577	0.93	0.97	0.97	0.89	0.94
	TEST	0.632	0.92	0.94	1.05	0.87	0.92
ANFIS	TRAIN	0.66	0.91	0.95	1.32	0.84	0.89
	TEST	0.742	0.82	0.85	1.39	0.8	0.83

Table 5 Performance of GPR and ANFIS models during 3 month lead time forecasting

Well location		Bellare			Guttigaru		
Statistical indices		RMSE (m)	NSE	CC	RMSE (m)	NSE	CC
GPR	TRAIN	0.74	0.89	0.91	1.105	0.86	0.91
	TEST	0.82	0.86	0.89	1.211	0.83	0.88
ANFIS	TRAIN	0.89	0.84	0.88	1.69	0.81	0.86
	TEST	1.09	0.79	0.84	1.847	0.76	0.82

Table 6 Performance of GPR and ANFIS models during 6 month lead time forecasting

Well location		Bellare			Guttigaru		
Statistical indices		RMSE (m)	NSE	CC	RMSE (m)	NSE	CC
GPR	TRAIN	0.96	0.83	0.87	1.225	0.83	0.85
	TEST	1.167	0.78	0.81	1.37	0.80	0.82
ANFIS	TRAIN	1.281	0.79	0.82	1.82	0.76	0.8
	TEST	1.41	0.74	0.78	1.97	0.71	0.75

Fig. 3 RMSE of GPR and ANFIS models at multistep lead time forecasting
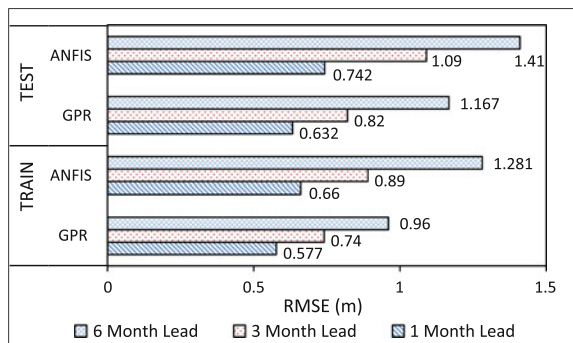
Figures 4 and 5 illustrate observed versus forecasted groundwater level time-series using GPR and ANFIS models. It can be seen from Figs. 4 and 5 that the GPR model can efficiently mimic observed groundwater level time-series better than ANFIS model during 1 month lead forecasting. Figures 6 and 7 are scatter
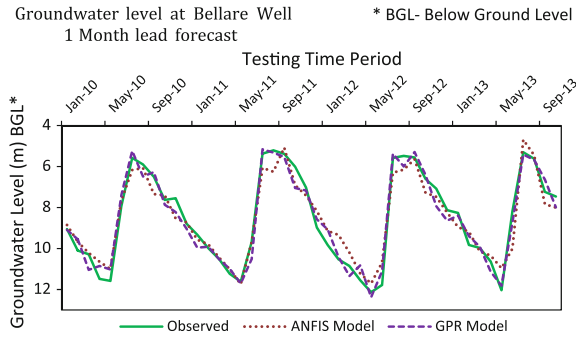
Fig. 4 Plot of observed versus forecasted groundwater level time-series with respect to the well location at Bellare of 1 month lead time forecasting models
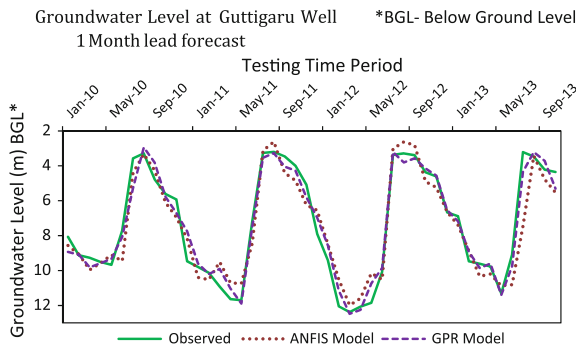

Fig. 5 Plot of observed versus forecasted groundwater level time-series with respect to well location at Guttigaru of 1 month lead time forecasting models
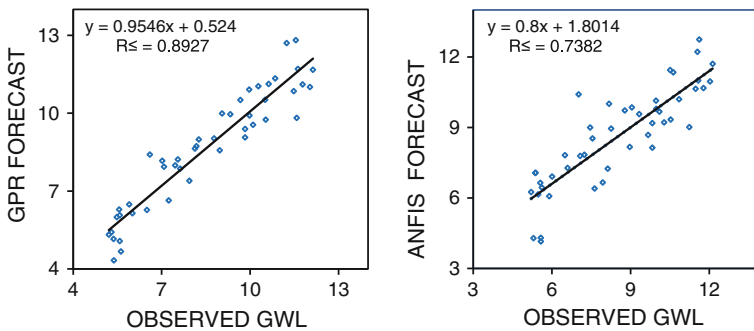

Fig. 6 Scatter plot of observed versus forecasted groundwater level with respect to well the location at Bellare of 1 month lead time forecasting models during test phase
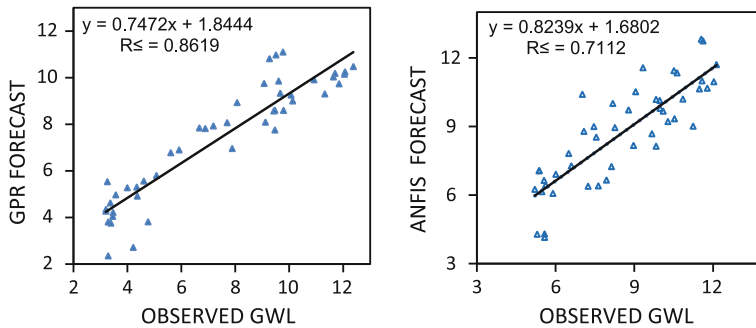
Fig. 7 Scatter plot of observed versus forecasted groundwater level with respect to well location at Guttigaru of 1 month lead time forecasting models during test phase

plots comparing the observed and forecasted groundwater levels using the GPR and ANFIS models for 1 month lead time forecasting during the testing period at the Bellare and Guttigaru sites. It can be observed that the band of scatter plot is very narrow and close to the line of perfect fit in case of GPR forecast, On the other hand ANFIS shows marginally lesser performance as compared to the GPR model in test phase. On a whole, it can be concluded that the GPR model provided more accurate forecasting results at both the study sites than the best ANFIS model at all the 1, 3 and 6 month lead times considered.

6 Conclusions

The application of the Gaussian Process Regression to forecast monthly groundwater level fluctuations at multistep lead times is investigated in the present study. ANFIS modeling is also adopted for comparative performance evaluation of the developed models. It is observed that the performance of the GPR is quite satisfactory providing relatively close agreement predictions when compared to that of ANFIS model in terms of the performance measures utilized in this study. It is envisaged that GPR model could serve as a better alternate for forecasting groundwater level fluctuation at multistep lead time. The GPR model has advantages over other models in terms of model accuracy, feature scaling, and probabilistic variance. In future one can test the applicability of GPR model with multivariate input data to forecast groundwater levels by including rainfall, temperature, and evaporation data.

References

1. Alley, W.M., Reilly, T.E., Franke, O.L.: Sustainability of Ground-Water Resources, p. 79. U. S. Geological Survey Circular 1186, Denver (1999)
2. Raghavendra, N.S., Deka, P.C.: Sustainable development and management of groundwater resources in mining affected areas. Procedia Earth Planet. Sci. **11**, 598–604 (2015). doi:10. 1016/j.proeps.2015.06.061
3. Taylor, C.J., Alley, W.M.: Ground-Water-Level Monitoring and the Importance of Long-Term Water-Level Data, p. 67. U.S. Geological Survey Circular 1217, Denver (2001)
4. Gupta, A.D., Onta, P.R.: Sustainable groundwater resources development. Hydrol. Sci. J. **42**, 565–582 (1997)
5. Adamowski, K., Hamory, T.: A stochastic systems model of groundwater level fluctuations. J. Hydrol. **62**, 129–141 (1983)
6. Ahn, H.: Modeling of groundwater heads based on second-order difference time series models. J. Hydrol. **234**, 82–94 (2000)
7. Bidwell, V.J.: Realistic forecasting of groundwater level, based on the eigenstructure of aquifer dynamics. Math. Comput. Simul. 12–20 (2005)
8. Sudheer, Ch., Mathur, S.: Groundwater level forecasting using SVM-PSO. Int. J. Hydrol. Sci. Technol. **2**, 202 (2012)
9. Daliakopoulos, I.N., Coulibaly, P., Tsanis, I.K.: Ground water level forecasting using artificial neural networks. J. Hydrol. **309**, 229–240 (2005)
10. Shirmohammadi, B., Vafakhah, M., Moosavi, V., Moghaddamnia, A.: Application of several data-driven techniques for predicting groundwater level. Water Resour. Manag. **27**, 419–432 (2013)
11. Nourani, V., Ejlali, R.G., Alami, M.T.: Spatiotemporal Groundwater Level Forecasting in Coastal Aquifers by Hybrid Artificial Neural Network-Geostatistics Model: A Case Study (2011)
12. Raghavendra, N.S., Deka, P.C.: Forecasting monthly groundwater table fluctuations in coastal aquifers using Support vector regression. In: Anadinni, S. (ed.) International Multi Conference on Innovations in Engineering and Technology (IMCIET-2014), pp. 61–69. Elsevier Science and Technology, Bangalore (2014)
13. Shiri, J., Kisi, O., Yoon, H., Lee, K.-K., Nazemi, A.H.: Predicting groundwater level fluctuations with meteorological effect implications-{A} comparative study among soft computing techniques. Comput. Geosci. **56**, 32–44 (2013)
14. Yoon, H., Jun, S.-C., Hyun, Y., Bae, G.-O., Lee, K.-K.: A comparative study of artificial neural networks and support vector machines for predicting groundwater levels in a coastal aquifer. J. Hydrol. **396**, 128–138 (2011)
15. Suryanarayana, C., Sudheer, C., Mahammood, V., Panigrahi, B.K.: An integrated wavelet-support vector machine for groundwater level prediction in Visakhapatnam, India. Neurocomputing **145**, 324–335 (2014)
16. Adamowski, J., Chan, H.F.: A wavelet neural network conjunction model for groundwater level forecasting. J. Hydrol. **407**, 28–40 (2011)
17. Maheswaran, R., Khosa, R.: Long term forecasting of groundwater levels with evidence of non-stationary and nonlinear characteristics. Comput. Geosci. **52**, 422–436 (2013)
18. Raghavendra, N.S., Deka, P.C.: Forecasting monthly groundwater level fluctuations in coastal aquifers using hybrid Wavelet packet—Support vector regression. Cogent Eng. **2**, 999414 (2015)
19. Brahim-Belhouari, S., Bermak, A.: Gaussian process for nonstationary time series prediction (2004)
20. Sun, A.Y., Wang, D., Xu, X.: Monthly streamflow forecasting using Gaussian process regression. J. Hydrol. **511**, 72–81 (2014)
21. Grbić, R., Kurtagić, D., Slišković, D.: Stream water temperature prediction based on Gaussian process regression. Expert Syst. Appl. **40**, 7407–7414 (2013)

22. Roberts, S., Osborne, M., Ebden, M., Reece, S., Gibson, N., Aigrain, S.: Gaussian processes for time-series modelling. Philos. Trans. A. Math. Phys. Eng. Sci. **371**, 20110550 (2013)
23. Yan, W., Qiu, H., Xue, Y.: Gaussian process for long-term time-series forecasting. In: Proceedings of the International Joint Conference on Neural Networks, pp. 3420–3427 (2009)
24. Box, G.E.P.: Non-Normality and tests on variances. Biometrika **40**, 318–335 (1953)
25. Rasmussen, C.E.: Gaussian processes in machine learning. In: Advanced Lectures on Machine Learning. Lecture Notes in Computer Science: Lecture Notes in Artificial Intelligence, pp. 63–71. Springer, Germany (2004)
26. Mackay, D.J.C.: Introduction to Gaussian processes. Neural Netw. Mach. Learn. **168**, 133–165 (1998)
27. Rasmussen, C.E., Williams, C.: Gaussian processes for machine learning. Adaptive Computation and Machine Learning, p. 272. The MIT Press, Cambridge (2006)
28. Jang, J.S.R.: ANFIS: adaptive-network-based fuzzy inference system. IEEE Trans. Syst. Man Cybern. **23**, 665–685 (1993)
29. Keskin, M.E., Taylan, D., Terzi, Ö.: Adaptive neural-based fuzzy inference system (ANFIS) approach for modelling hydrological time series. Hydrol. Sci. J. **51**, 588–598 (2006)
30. Hall, M., National, H., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., Witten, I.H.: The WEKA data mining software: an update. SIGKDD Explor. **11**, 10–18 (2009)