GROUNDWATER LEVEL FORECASTING MODEL USING HYBRID SUPPORT VECTOR REGRESSION -PARTICLE SWARM OPTIMIZATION FOR AQUIFER IN UDUPI REGION

Supreetha B.S*
Faculty of Electronics and Communication Engineering
Manipal Institute of Technology
Manipal Academy of Higher Education, Manipal, India

Narayan Shenoy
Faculty of Civil Engineering
Manipal Institute of Technology
Manipal Academy of Higher Education, Manipal, India

Prabhakar Nayak
Faculty of Electronics and Communication Engineering
Manipal Institute of Technology
Manipal Academy of Higher Education, Manipal, India

*Corresponding author Email: supreetha.bs@manipal.edu

ABSTRACT

Groundwater level in Udupi region decline invariably because of variation in rainfall and the ruggedness of topography and porous nature of lateritic rock. The groundwater level forecasting model needs to be developed for investigating the varying water level to save this precious water resource. This research work is focused in the Udupi region located in Karnataka state, India for two different types of major geological formations, lateritic terrain and Banded Gneissic Complex (BGC). In this paper hybrid Particle Swarm Optimization guided support Vector Regression (SVR) approach is employed to forecast the future trend. The particle swarm optimization algorithm (PSO) is used to select optimal SVR based parameters. Hybrid SVR –PSO model is tested with historical groundwater level data and rainfall data collected from Udupi Region. Forecasting performance of the Artificial Neural Network (ANN) and SVR models were analyzed using statistical metrics like MAE, NMSE. The ANN model shows high performance for larger dataset and SVR models shows high performance for limited dataset. The result indicates that the SVR shows less relative error than ANN and SVR could be a better alternative for forecasting
Groundwater has always kept its supreme position as one of the cardinal resources to provide drinking and agriculture water requirements. In several regions of the cosmos, groundwater plays its pivotal role as a noteworthy water resource for various needs [1]. In view of these factors, groundwater has become a significant and dependable resource in supplying consumption requirements of varied clients. Groundwater reservoir, in essence, represents an intricate system which encounters both natural and artificial constraints, bringing hassles on the entire system of aquifer in various chronological levels leading to avoidable variations in groundwater level. In watersheds where data is restricted and achieving precise predictions is rather more significant than comprehending the basic mechanisms, data based models surface as another viable substitute [2].

For analyzing the depletion and variation of water table due to the unsustainability of the environment monitoring and collecting the groundwater level data is very much required [3]. The numerical and statistical forecasting methods were generally used for this task. Numerical methods are accurate in calculations but not suitable for accurate forecasting due to the irregular variation in the patterns of data, therefore forecasting by this approach was not so successful [4]. Data driven techniques are better used for developing groundwater level forecasting models using quantitative historical data with data driven algorithms.

2. RELATED WORK

A large variety of groundwater level forecasting models were proposed in the literature for limited dataset [5]. Data driven time series models are commonly used approach have received much interest in the recent literature.

The PSO is a nature inspired algorithm which mimics the collective flocking of birds. In Particle Swarm Optimization algorithm, the solution called particles are found by updating velocity and position of solutions and finding optimal solution [6].

A SVM and ANN based groundwater forecasting model is proposed by Heesung et.al [7]. The model compared two data driven prediction models for small dataset in aquifers of coastal region. Model performance implicated that SVM was better than ANN and generalized well with lead times compared to ANN model. Behzad et.al [8] studied Support Vector Machines for forecasting hydrological parameters with different weather conditions and pumping rate. They found that SVM gives good accurate results with limited data, whereas ANN requires large amount of data and there is no consistency in the results. Therefore, they suggested that, for real time analysis of hydrological parameters, SVM may be more suitable. The hybrid model of Wavelet Packet–Support Vector Regression was proposed by Sujay Raghavendra and Pareshchandra Dekha [9] to predict levels of identified wells from Dakshina Kannada district. Discrete wavelet packet transformation was used to
preprocessing of data and hybridized used support vector regression for predicting the levels of groundwater. They used WEKA tool to simulate and estimate parameters related to SVR. The efficiency of SVR based approach relied on selecting ideal values, hyper parameters related to kernel function, loss function and regularization parameters.

Sudheer et.al [10] recommended support vector machine with Particle Swarm Optimization (PSO) for predicting surface water flow. They used Particle Swarm Optimization algorithm with support vector machine to increase the prediction accuracy by analyzing SVM parameters with PSO.

Gong et.al [11] developed 3 models Neural Network model, Support Vector Machine and Fuzzy Inference system. They applied the developed models for forecasting the level of groundwater for two identified wells, located adjacent to a lake in USA. Level for 2 wells near lake in Florida, USA. They used datasets collected for a period of one decade. The performance of these models are evaluated using statistical performance measures NMSE, Correlation Coefficient and NS. The model effectively proves the applicability of all these three soft computing models in groundwater level forecasting, considering the fluctuations in lake level.

3. STUDY AREA AND DATA

In this work, the monthly groundwater level and rainfall data of Udupi region were used in evaluating the performance of SVM-PSO hybrid approach. The location map of Udupi region and observation well were as shown in Fig.1.

![Figure 1](image)

*Figure 1 Location map of Udupi region and observation well*

The systematic work on geology, geomorphology of certain parts of Udupi district were conducted by early researchers such as Radhakrishna Vaidhyanathan and Balasubramanyana (1994) [12]. Based on their studies laterites, gneisses granites, dolerite dykes and coastal sediments rock by Banded Gneissic Complex (BGC) and laterites.

The gneisses of Udupi show all the major characteristics of older gneissic complex which are common in the other parts of the craton. They are layered and banded complex consisting of quartzo felspathic biotitic gneissic rocks exposed over large parts of the Udupi taluk as small hillocks. Udupi district consists of basement peninsular gneisses exposed in the major portion of the district. From the ground water point of view, the gneisses are classified as
crystalline formations. The fracture system developed along with joints and faults traversing the rocks facilitate groundwater circulations and hold moderate quantity of water. Groundwater generally occurs in the water table conditions in the weathered mantel and also under semi confined conditions in the deeper fractures.

Another major type of geological formation in Udupi district is lateritic terrain. Two major types of laterites are identified in Udupi Taluk. They are primary laterites and detrital laterites. Primary laterites developed in situ over crystalline rocks viz., gneisses. Primary laterites are more homogeneous, less sandy and are developed in situ through alteration of the underlying rocks. Detrital laterites are derived from reworking of primary laterites. They generally contain sands and rounded pebbles of quartz [13].

4. PROPOSED METHODOLOGY

The proposed research work implements a groundwater level forecasting using hybrid soft computing approach which consists of hybrid PSO based Support Vector Regression with RBF kernel as depicted in Fig.2. The accuracy of the SVR based groundwater level forecasting model depends on tuning of SVR parameters. In this work Swarm Intelligence (SI) based technique called PSO is used to select the ideal parameters for Support Vector regression.

![Figure 2 Schematic diagram depicting PSO-SVR methodology](image)

5. INTRODUCTION TO SVM AND ITS PARAMETERS

Optimization problem can be solved using SVR method, considering \{(x_1, y_1), ...,(x_m, y_m)\} for the datasets where each \( x_i \in R^n \) which resembles the sample space for the input and target value can be considered as \( y_i \in R \) for \( i=1,2,\ldots,m \), whereas \( m \) represents size dataset to be trained.
The above problem can be formulated as a Minimization problem given as:

\[
\frac{1}{2}\|w\|^2 + C \sum_{i=0}^{m} (\xi_i + \xi_i^*)
\]

Subject to

\[
\begin{align*}
\gamma_i - \langle w, x_i \rangle & - b \leq \varepsilon_i + \xi_i \\
\langle w, x_i \rangle + b - \gamma_i & \leq \varepsilon_i + \xi_i^* \\
\xi_i, \xi_i^* & \geq 0 \quad i = 1, ..., m
\end{align*}
\]

Where the function \( \Theta \) is used for mapping input to a higher dimensional space. Where \( \xi_i \) and \( \xi_i^* \) represents the upper and the lower training error to the \( \varepsilon \) insensitive tube \( \gamma_i - \langle w, x_i \rangle - b \leq \varepsilon \).

The parameters related to SVR are the cost of error \( C \), the width of the tube \( \varepsilon \) and the mapping function \( \Theta \). The constraint of increased data tendency \( x_i \) can added within the tube with less error by minimizing the objective function. The under fitting and overfitting of the training data by reducing the training error and the regularization term can be avoided using SVR technique. Hence for regression problem SVM seems to be more accurate and flexible method.

The Radial Basis Function kernel used is given by:

\[
k(x_i, x) = \exp\{-\gamma |x - x_i|^2\}
\]

Where \( \gamma \) represents the kernel width. The value of \( \gamma \) lies between 0 and 1.

The SVM model has 3 parameters represented \( C \), \( \varepsilon \) and \( \gamma \). Global optimization algorithm PSO is helpful in determining the parameters of SVM. In Particle Swarm Optimization algorithm, the solution called particles is found by exploring through the problem space specified by the current optimum particles. The upper and lower bounds of \( C \), \( \varepsilon \) and \( \gamma \) are initialized and the values of these parameters are set and the fitness is evaluated. The normalized mean square error (NMSE) serves as the fitness criteria for identifying the suitable SVM parameters. The mathematical equations used to formulate Particle Swarm Optimization are given by:

\[
\begin{align*}
V_{id} &= W^0 V_{id} + C_1 \text{rand}(P_{id} - X_{id}) + C_2 \text{rand}(P_{gd} - X_{id}) \\
X_{id} &= X_{id} + V_{id}
\end{align*}
\]

Where \( C_1 \) and \( C_2 \) are the acceleration coefficients and \( \text{rand} \) are random numbers between 0 to 1 and \( W^0 \) is called inertia weight to control the current velocity.

6. RESULTS AND DISCUSSION

6.1. Spatial variability

The collected ground water level data from Udupi region were used to develop groundwater level map. Groundwater level data of 30 located points from District Commissioner office, Udupi district for the year 2016 and 2017 were used for mapping using QGIS software. The depth to groundwater level for pre-monsoon and post monsoon period for the year 2016 and 2017 is as shown in Fig.3.
During premonsoon depth to water level varies between 7 to 9 mbgl and post monsoon water level varies between 5 to 7 mbgl in the Udupi district. The groundwater level is deepest represented by blue color before commencement of monsoon and shallowest represented by red color in the map for post monsoon period. The rise in water level after monsoon indicates the building up of groundwater storage, which gets depleted during non-monsoon period.

6.2. Seasonal groundwater level forecasting
The major part of Udupi region is covered by gneisses and laterites terrain types. The developed forecasting model was tested over the historical groundwater level data and rainfall for a period of 4 years collected from Udupi region for gneissic and lateritic terrain formations.

6.2.1. Lateritic terrain
The yearly forecasted groundwater level using SVM-PSO model for lateritic terrain during testing period up to 2019 and 2022 are shown in Fig.4

Figure 3 Groundwater level map for pre and post monsoon for 2016 and 2017

Figure 4 Yearly forecasted groundwater level for lateritic terrain
The seasonal forecasted groundwater level using SVM-PSO model for pre monsoon period is as shown in Fig.5 below.

![Figure 5](image-url-for-premonsoon)  
**Figure 5** Forecasted groundwater level for the pre monsoon period for lateritic terrain

The seasonal forecasted groundwater level using SVM-PSO model for post monsoon period is as shown in Fig.6 below.

![Figure 6](image-url-for-postmonsoon)  
**Figure 6** Forecasted groundwater level for the post monsoon period for lateritic terrain

6.2.2. Banded Gneissic Complex

The seasonal forecasted groundwater level for post monsoon and premonsoon period using SVM-PSO model for Banded Gneissic Complex terrain during testing period up to 2020 is as shown in Fig.7 and Fig.8 below.
Figure 7 Forecasted groundwater level for the post monsoon period for BGC

It is observed that support vector regression based forecasting model is able to forecast the groundwater level of identified open well maximum up to 4 years ahead. The forecasted groundwater level in lateritic terrain varies rapidly compared to BGC. The forecasted groundwater level of 2020 will be declined up to 13 to 14m in a well with depth 16m in dry season and after 2021 the GWL will be recovering. The seasonal groundwater level variation in BGC is less compared to lateritic terrain.

Gneissic granites are hard rocks from the groundwater occurrence point of view and are devoid of primary porosity. The occurrence, distribution and movement of groundwater depends on the presence of surface and deep-seated fractures and joints also on the extent of weathering. However due to intensive weathering and presence of surface and subsurface joints and fractures these formations are found to yield good quantity of groundwater.

6.2.3. Comparative analysis

The relative MAE is more stable relative error metric used to compare two forecasting models. The ratio of MAE measure obtained by SVR model and ANN model is given by

\[ \text{RelMAE} = \frac{\text{MAE}_{SVR}}{\text{MAE}_{ANN}} = 0.96 \]

The value near 1 indicates that the magnitude of the two errors is approximately equal. The above results show that the SVR shows less relative error than ANN.
The NMSE plot for the year 2019 and 2020 is as shown in Fig.9 below.

![Image](image_url)

**Figure 9** Comparative Plot

SVM-PSO hybrid model gives accurate stable results compared to ANN for smaller datasets.

![Image](image_url)

**Figure 10** RMSE plot

7. CONCLUSIONS
The two different geological formations, lateritic terrain and Banded Gneissic Complex found in Udupi region were investigated to forecast the groundwater level. The rainfall and groundwater level data were collected from the observation wells identified in Udupi Region. The model was implemented using Support Vector Regression based approach and Particle Swarm Optimization was used to find optimal parameter of SVR and Kernel. It is recommended that SVR-PSO based hybrid approach for forecasting groundwater level with limited data. The flow of groundwater is unique for different geological formations. In hard rock terrains GW potential is relatively complex due to highly variable nature of geological terrain. These formations are found to be productive in terms of ground water due to intensive weathering and presence of surface and subsurface joints and fractures. We observed rapid seasonal variation of groundwater level in lateritic terrain compared to BGC. The laterites exhibit variable permeability due to highly heterogeneous nature of this material. The laterites of Udupi region are porous and possess medium to high permeability. The high porosity and permeability favors the movement of groundwater. Though the highly porous and permeable nature of laterites are favorable for groundwater storage, the ruggedness of topography and porous nature of this rock induces the infiltrated water to escape the base flow. This is one of the reason for drying up of most of the open wells during summer months in Udupi region.
The hybrid support vector regression based forecasting model were compared with ANN model. The results with SVR-PSO hybrid model show more accurate and stable compared to ANN model with limited data.

ACKNOWLEDGEMENT
The authors would like to thank statistical department and geology department, District Commissioner office Udupi District for providing valuable rainfall data, Groundwater level data required for research and Manipal Institute of Technology Karnataka for the necessary infrastructural support.

REFERENCES


