JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 1895-1904
https://publishoa.com
ISSN: 1309-3452

Analysis of Groundwater Level Fluctuations and its Association with Rainfall Using Statistical Methods

V Anantha Natarajan[1*], V Tamizhazhagan[2], Naresh Tangudu[3], M Sunil Kumar[4]

[1,4]Department of Computer Science & Engineering, Sree Vidyanikethan Engineering College, AP, India.

[2]Department of Information Technology, Annamalai University, India.

[3] Department of IT, Aditya Institute of Management and Technology, India.

E-mail: vananthanatarajan@vidyanikethan.edu[1], rvtamizh@gmail.com[2], itsajs@gmail.com[3], sunilmalchi1@gmail.com[4]

Received 2022 April 02**; Revised** 2022 May 20**; Accepted** 2022 June 18**.**

ABSTRACT

The state of Andhra Pradesh accounts for 5.3% of the nation's net groundwater irrigated area, with roughly half of all irrigation dependent on groundwater sources. With paddy making up about 70% of the state's total irrigated land, the state continues to be one of the top exporters of rice, but groundwater depletion poses serious problems to not only agricultural productivity and rural lifestyles, but also to the country's food security. In order to pinpoint the time when groundwater resources are being overused, the research investigates the relationship between groundwater level and rainfall using overly-simplistic assumptions about aggregate abstraction and groundwater recharge estimates. Inter-annual variations in groundwater level are closely related to inter-annual variations in rainfall and exploitation of groundwater resources for irrigation. To investigate the trends of changes in the groundwater level, trend analysis, regression analysis, and correlation analysis were carried out.

Keywords-

INTRODUCTION

Water that is found underground in saturated regions beneath the surface of the earth is known as groundwater. The water table is the term for the saturated zone's upper surface. One of the most significant natural resources for the country is groundwater, which is found in aquifers beneath the Earth's surface. The majority of the water that county and city water departments distribute to residences and businesses comes from groundwater. The majority of the rural population, who do not receive water delivery from a government water department or private sellers, rely on the ground water source for drinking and other purpose. According to the 2018 UNESCO World Water Development Report, India is the country that extracts the most groundwater globally [2]. According to the CGWB, many areas of India are experiencing a growing groundwater shortage because to the 230 billion meter cubes of groundwater that are extracted each year to irrigate agricultural lands. In India, there has been an estimated loss of 122-199 billion meter cubes of groundwater.

One of the most significant agrarian states in the nation is Andhra Pradesh. Nearly 37% of the state's GDP comes from agriculture. The majority of the population—nearly 70%—depends on agriculture for a living [3]. The state of Andhra Pradesh accounts for 5.3% of the nation's net groundwater irrigated land and uses groundwater for 49% of all irrigation. Today, excessive groundwater use and depletion pose severe threats to the nation's food security as well as agricultural production and rural lives. India has a 68 percent drought-prone area and a 12 percent flood-prone area. The harsh climate conditions generally hinder agricultural productivity. The "climate refugee," a special type of factor that could become invasive in cultivation and have an indirect impact on gross domestic products, may be brought about by climate change (GDP). The surface water storage is primarily impacted by the geographical and temporal fluctuations in rainfall pattern. Groundwater development efforts for irrigation, residential use, and industry are prompted by the shortfall. The abundant aquifers have tremendous groundwater storage, and these are the only naturally regenerating resources. The main source of replenishment is large recharge from rainwater percolation, although there is also some input from canal leakage and irrigation return flows.

STUDY REGION

Andhra Pradesh State, with a land area of 1,63,000 km2, is the seventh-largest state in India. It is located between EL 76° 45' and 84° 47' and NL 12° 37' and 19° 09'. State borders with Telangana, Chattisgarh, and Orissa states to the north, Tamil Nadu and Karnataka to the south, Karnataka to the west, and the Bay of Bengal to the east (970 km). The state's tropical climate is affected by geographical changes as well as maritime impact. In comparison to the coastal zone, the Deccan Plateau enjoys a more temperate temperature. During the southwestern monsoon, Vishakhapatnam and its surrounding area's Eastern Ghats, which function as a barrier to easterly winds in conjunction with depression from the Bay of Bengal, play a key role [4].

Annual Rainfall and Groundwater Statistics

Due to heavy rainfall this year, the State's groundwater level has increased. The average groundwater level in the state rose by 0.73 meters below ground level (MBGL) in January compared to last year. In January 2021 the average ground water depth was 7.24 MBGL and this year January it was 6.51 MBGL. Due to record high rainfall in the district of Anantapur the ground water level has increased considerable high when compared to other districts. In the coastal region of the state of Andhra Pradesh the average ground water level improved from 7.45 MBGL last year January to 6.79 meters this year January (0.66 metres increased). In contrast the ground water level of Rayalaseema region improved by 0.89 meters when compared to last year (Jan 2021 – 6.77 MBGL and Jan 2022 – 5.88 MBGL) [5].

District wise increase in ground water level due to heavy rainfall in the beginning of this year is summarized in Table 1. Despite of heavy rainfall in the state, the ground water level in three districts has reduced (Chittoor decreased by 0.91 meters, Kurnool 0.63 meters, and east Godavari 0.39 meters) when compared to last year. The expansion of land under cultivation and the sinking of additional bore wells to collect water are thought to be the primary causes of the decrease in groundwater level in those three regions.

Table 1 Summary of Ground Water Level in AP

S No	Region	Ground Water Level in MBGL		
		Jan 20	Jan 21	Jan 22
1	Anantapur	16.8	11.98	7.09
2	Vizianagaram	5.43	6.01	4.26
3	Srikakulam	4.20	5.69	4.49
4	Kadapa	12.10	4.94	3.84
5	Chittoor	19.32	6.13	7.05
6	Kurnool	7.29	6.66	7.07
7	East Godavari	9.29	8.90	8.98
8	West Godavari	20.31	17.86	18.64
9	Coastal AP	9.34	7.11	6.79
10	Rayalaseema	15.17	7.04	5.88

During the summer, Anantapur's groundwater table drops to the record lows in the state—roughly 30 MBGL. The levels have decreased to 11.98 metres below ground when compared to the ground water level in January 20 (16.8 MBGL), though. When comparing Jan 20 and Jan 21, Srikakulam and Vizianagaram saw a reversal in trend because these two districts had received insufficient monsoon rains. In the year 2017 the ground water level at Vempalle mandal of Kadapa district was 114.01 MBGL which is the deeper water table in the state. As borewells have been drilled to a depth of more than 1,000 feet to access the third aquifer system, the water level is rapidly decreasing in that region. In addition to overuse of the water resource, the poor recharging due to the area's geology is another factor contributing to the low groundwater table. The ground water will typically be on top when shale is below and limestone is above.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 1895-1904
https://publishoa.com
ISSN: 1309-3452

However, the limestone layers in that area are buried beneath the shale, and as a result, the water is only accessible at deeper level.

The above data presents two contrasting scenarios; i. due to heavy rainfall in the beginning of the year 2022 the ground water level improved throughout the state except srikakulam, kurnool and east Godavari. The water level went slightly deeper due to the exploitation of ground water for irrigation and other purpose. ii. Due to scarce rainfall in the year 2021 the Srikakulam and Vizianagaram witnessed a decrease in ground water level by 0.98 and 0.58 meters respectively. In general it can be inferred that if the annual rainfall in the region is high then correspondingly the ground water level also increases but at times due to over exploitation of ground water resources there may be reversal trend witnessed in the region.

In India, the four monsoon months of June through September account for nearly 75% of total rainfall (shown in Fig.1). The extreme hydrological events (widespread drought and floods) caused by the yearly variations in monsoon rains have a serious negative impact on groundwater levels, agricultural yield, population, and the national economy. The annual rainfall, and moisture level in the soil directly influences the droughts, floods, and desertification.

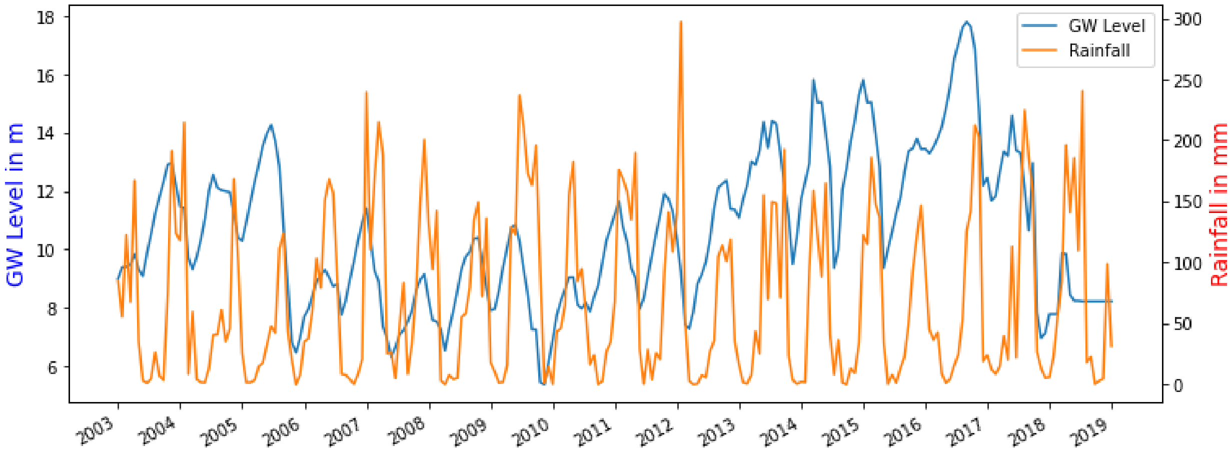

Fig. 1 Comparison of Rainfall and Groundwater over the period 2003 -2021

The division of a time series into systematic and non-systematic components is a helpful abstraction for time series analysis. Time series' systematic components can be characterized and modelled because they exhibit consistency or repetition [6]. The time series' non-systematic components cannot be directly modelled. Three systematic components—level, trend, and seasonality—along with one non-systematic component noise are believed to make up any given time series. The following is a definition of these elements:

Level : The series' median value.

Trend : The value's upward or downward movement over time.

Seasonality : The series' recurring brief cycle.

Noise : The series' random variation.

In the first step of a traditional decomposition, the trend-cycle is estimated using the moving average approach [7]. The moving average with an order m can be expressed as

$$\hat{T}_t = \frac{1}{m}\sum_{j=-k}^{k} y_{t+j} \qquad \text{Eq. 1}$$

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 1895-1904
https://publishoa.com
ISSN: 1309-3452

where $m = 2k + 1$. Values from the time series within k periods of t are averaged to produce an approximation of the trend-cycle at time t. The averaging reduces some of the randomness in the time series, providing a smoother trend-cycle component. The trend derived from the rainfall and groundwater time series is presented in Fig. 2.

Fig. 2 Comparison of trends in Groundwater level and Rainfall

NON-PARAMETRIC METHODS BASED ANALYSIS

The detection of trends in many hydrologic series, such as rainfall, humidity, water evaporation, and radiation, is frequently done using non-parametric approaches. In this study, the water table and rainfall patterns in the study region were identified using the Mann-Kendall (MK) approach [8]. The statistic known as Kendall's coefficient is used to determine the ordinal relationship between two measured items. It measures the Kendall rank correlation coefficient. The coefficient-based test is a non-parametric hypothesis test to analyze the statistical dependence with respect to the τ coefficient. The Kendall correlation between two variables will be higher when observations rate the two variables similarly and lower when observations rank the two variables differently.

Let us consider that a set of observations represented as $(x_1, y_1), \ldots\ldots (x_n, y_n)$ derived from two different random variables denoted by X and Y. A pair of observations denoted by (x_i, y_i) and (x_j, y_j) can be concordant if $x_i > x_j$ and $y_i > y_j$ holds true or $x_i < x_j$ and $y_i < y_j$ holds true; else the pair is said to be discordant. Based upon this assumption the Kendall τ coefficient is expressed mathematically as:

$$\tau = 2 * \frac{(number\ of\ concordant\ pairs) - (number\ of\ discordant\ pairs)}{n(n-1)} \qquad \text{Eq. 2}$$

The denominator represents the number of pair combinations that can be derived from the two random variables and it makes the coefficient to lie between the range $-1 \leq \tau \leq 1$. It begins by presuming that there is no trend (i.e., tau () is assumed to have a value of 0). If a data value from a later period is greater than a data value from an earlier time when comparing two consecutive years, τ is fixed as 1. Similar to this, is reduced by 1 if a data value from a later period is less than a data value from a previous time. As a result, a high positive number indicates an upward trend, whereas a low negative value indicates a downward trend. The magnitude of the observed trend can be estimated using the Sen's slope estimation approach represented as;

$$T_i = \frac{x_j - x_k}{j - k}\ for\ i = 1,2,3, \ldots \ldots \ldots, N \qquad \text{Eq. 3}$$

where x_j and x_k are values of the series at two different instances and the sen's slope estimator denoted as β can be calculated by finding the median of N values of T_i [9]. The positive and negative value of the slope denotes the upward and downward trend.

Table 2. Summary of Trend analysis using Mann-Kendall test

Month	Tau	Sen's slope	p value	Average Ground Water in MBGL	Average Rainfall in mm
January	0.1636	0.0769	0.3143	9.06	6.48
February	-0.3113	-0.1285	0.0555	9.64	6.95
March	0.0421	0.0260	0.8227	10.31	11.82
April	0.0632	0.0150	0.7246	10.88	18.71
May	-0.3368	-0.0471	0.0398	11.38	44.65
June	0.2053	0.0178	0.1939	11.92	90.78
July	-0.1108	-0.0045	0.4954	12.01	118.01
August	-0.2269	-0.0178	0.1628	11.77	137.99
September	-0.0369	-0.0035	0.8202	11.24	142.15
October	-0.0686	-0.0046	0.6730	10.38	139.29
November	-0.1693	-0.0052	0.2987	9.56	89.99
December	0.0792	0.0127	0.6263	8.89	24.74
Total Rainfall	0.1166	0.0069	0.0070	10.59	843.48
SW Monsoon	-0.0416	-0.0026	0.5819	11.74	487.39
NE Monsoon	0.1013	0.0045	0.2535	9.61	254.03
Winter Season	-0.0706	-0.0334	0.5215	9.35	13.44
Summer season	-0.0039	-0.0001	0.9643	10.86	75.18

GRADIENT BOOSTING REGRESSION

A statistical method for assessing relationships between a crucial response variable and a variety of other factors is regression analysis. This approach is used to identify the independent variables that are connected to the dependent variable and to investigate the nature of these connections. Prediction and estimation are two common uses of regression analysis. A linear regression model can be expressed mathematically as follows;

$$y_i = \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \cdots \dots \dots + \beta_m x_{im} + \varepsilon_i \; i = 1,2,\dots,n \qquad \text{Eq. 4}$$

The above mathematical expression relates m independent variables to a dependent variable y, and the ε is a random error. The $cov(\varepsilon_i, \varepsilon_j) = 0$ and $\beta_0, \beta_1, \dots., \beta_m$ denotes the coefficients or the parameters of the regression model. Boosting algorithms are built on additive modelling. The concept is straightforward: combine a number of smaller expressions to create a complicated function. Gradient Boosting is the process of combining several simpler models to create a more complicated final model. Gradient boosting uses a weighted sum of the base learners to learn a model [10].

Consider a training dataset denoted as $\{(x_i, y_i)\}_{i=1}^{n}$ and objective function $L(y, F(x))$. The following steps describes the M step iterative gradient boosting regression algorithm.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 1895-1904
https://publishoa.com
ISSN: 1309-3452

Algorithm:

Step 1: Assign a constant value to the model initially:

$$F_0(x) = arg\min_{\gamma} \sum_{i=1}^{n} L(y_i, \gamma) \qquad \text{Eq. 5}$$

Step 2: Implement the following statements iteratively $m = 1\ to\ M$

i. Estimate the pseudo –residuals:

$$r_{im} = -\left[\frac{\partial L\big(y_i, F(x_i)\big)}{\partial F(x_i)}\right]_{F(x)=F_{m-1}(x)} for\ i = 1, \dots., n \qquad \text{Eq. 6}$$

ii. Train a weak learner usually a decision tree, $h_m(x)$ using the training set $\{(x_i, r_{im})\}_{i=1}^{n}$

iii. Solve the following optimization problem to find the value of γ_m

$$\gamma_m = \operatorname{argmin}_{\gamma} \sum_{i=1}^{n} L\big(y_i, F_{m-1}(x_i) + \gamma h_m(x_i)\big) \qquad \text{Eq. 7}$$

iv. Update the $F_m(x)$:

$$F_m(x) = F_{m-1}(x) + \gamma_x h_m(x) \qquad \text{Eq. 8}$$

RESULTS AND DISCUSSION

The frequency and severity of heavy monsoonal rain events in the area had seen significant modifications. According to the analysis of rainfall, Varanasi had an average annual rainfall of 843.48mm from 2003 to 2022 (presented in Fig. 3). Following the SW monsoon, winter and summer, rainfall showed declining patterns, while after the SW monsoon (NE monsoon), rainfall showed increasing tendencies (Table 2). The NE monsoon supplied 30.5 percent of the yearly rainfall, while the SW monsoon contributed 58.8 percent with a little decreasing tendency. The annual rainfall over the state may have contributed to the declining trend in total precipitation. Rainfall during the monsoon season made up 89.3% of the total yearly rainfall.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 1895-1904
https://publishoa.com
ISSN: 1309-3452

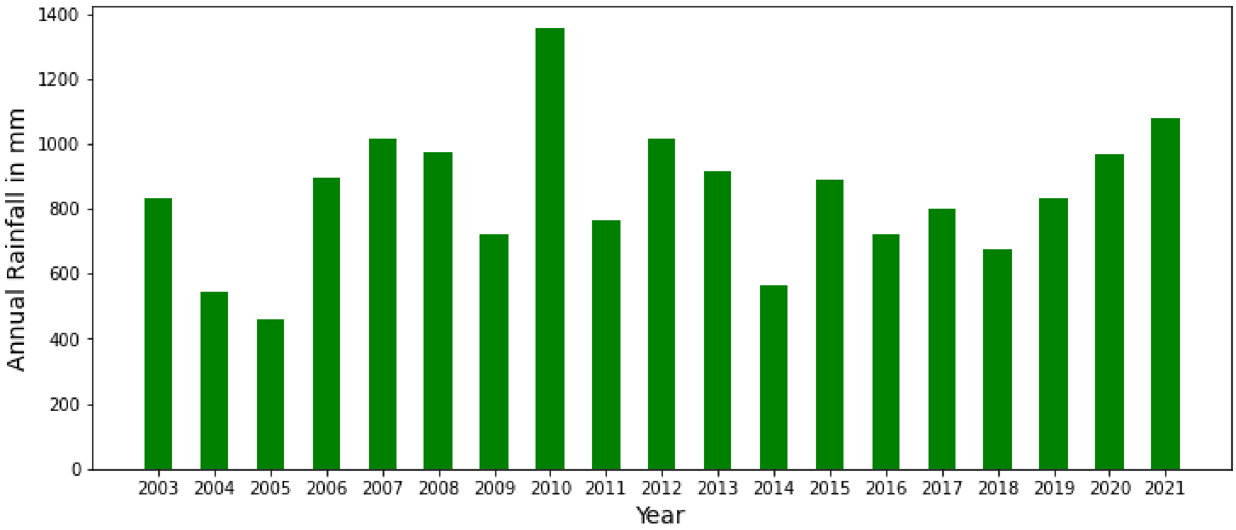

Fig. 3 Comparison of Annual Rainfall in mm from 2003 to 2002

The ground water level was observed to be 11.73 MBGL on an average during the SW monsoonal period and average level of 9.61 during the NE monsoonal period. During the winter and summer season the ground water level was observed to be 9.35 and 10.85 MBGL respectively. Due to heavy rainfall during the SW monsoonal period the ground water recharge is quite high and the ground water level increases considerably. The rabi season due to scarce rainfall in many parts of the state the cultivated lands are majorly irrigated from ground water sources and hence the respective period the ground water levels are observed to be low. As per the report published by Ministry of Jal Shakti on March 2022, about 21 % of the bore wells across the state has observed a fall in the ground water level and among 556 well were the ground water level has increased; nearly in 7% of the well the water level has increased more than 4 meters (shown in Table 3) [11].

Table 3. Status of groundwater level in wells as of March, 2022

No. of wells Analyzed	Rise						Fall						Rise		Fall		Wells showing no change	
	0-2 m		2-4 m		>4 m		0-2 m		2-4 m		>4 m							
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
706	419	59.3	87	12.3	50	7.1	124	17.6	14	2	11	1.6	556	79	149	21	1	0

The ground water recharge estimations presented in Fig. shows that the recharge pattern closely matches the annual rainfall distribution in the state. Some empirical relationships have been constructed for computing the natural recharge to ground water from rainfall based on studies conducted by various scientists and organizations regarding the association of ground water level fluctuation and rainfall. One such empirical relationship derived to estimate the ground water recharge with respect to the annual rainfall (applicable in areas where annual rainfall exceeds 400 mm) can be mathematically expressed as $R = 2.0\,(P - 15)^{0.4}$ where R denotes the net annual recharge in inches and P denotes the annual rainfall in inches [12]. Later the base representation was modified by few researchers at IIT Roorkee as follows; $R = 1.35(P - 14)^{0.5}$ [13].

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 3, 2022, p. 1895-1904
https://publishoa.com
ISSN: 1309-3452

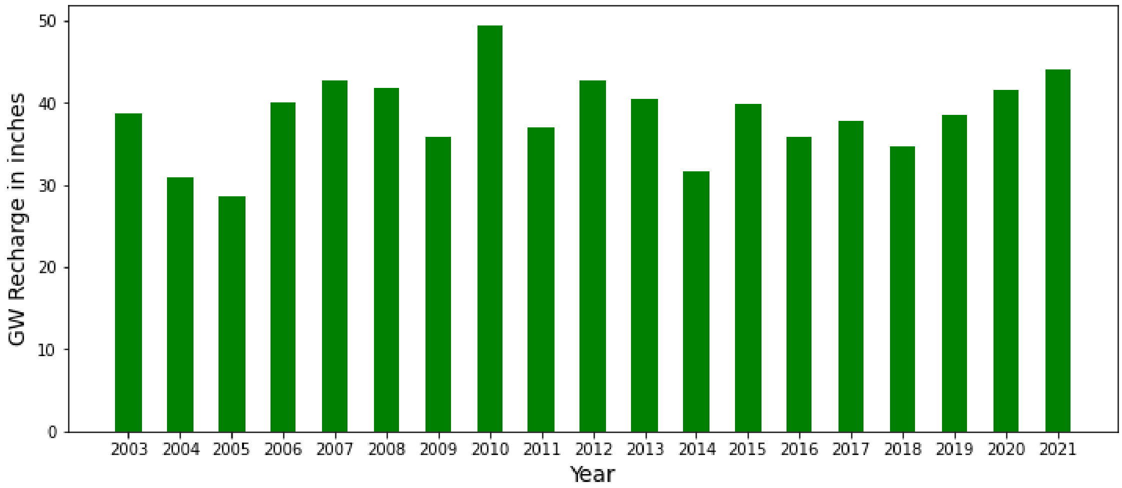

Fig. 4 Estimated Groundwater for the period from 2003 - 2021

Regression based analysis

Groundwater resources include the recharging of rainfall to the ground and groundwater evaporation. The primary source of groundwater recharge is precipitation. Fig. 2 illustrates analyses of the impact of precipitation on changes in the groundwater table. The annual rainfall and use of groundwater for irrigation have a direct impact on the diachronic change of the groundwater table. The groundwater table typically increases with precipitation and decreases with increased or excessive resource extraction. The comparison of the groundwater level observed and predicted using the regression analysis is presented in Fig. 5

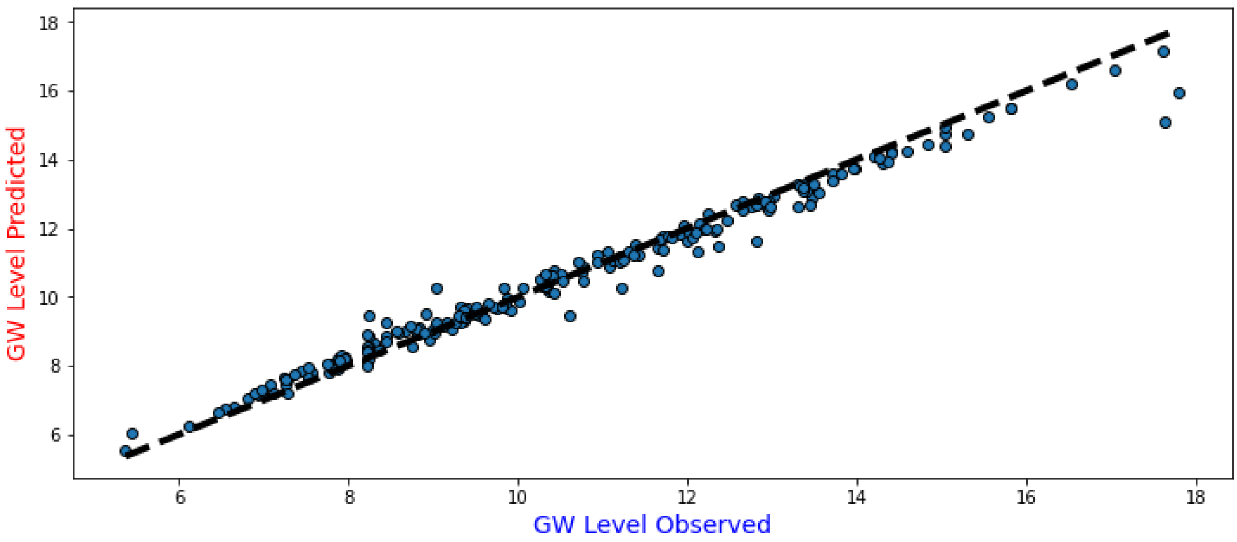

Fig. 5 Comparison of Observed and Predicted Groundwater Level in MBGL

Fig. 2 displays a typical groundwater and monthly rainfall trend. The fluctuation cycle is timed with rainfall events, and the groundwater trend curve displays many peaks and multiple valleys. There is a perfect 1:1 correlation between each curve peak of the groundwater level and each curve peak of rainfall. The groundwater level responds quickly to changes in rainfall. Less rain fell between the years of 2004 and 2006, and there were negligible changes in the level of the groundwater. The groundwater level showed a declining trend during a dry spell due to groundwater pumping for agriculture, and it continued to decline until the next monsoon season. The frequent monsoonal rainstorms from June to December caused major changes in groundwater level.

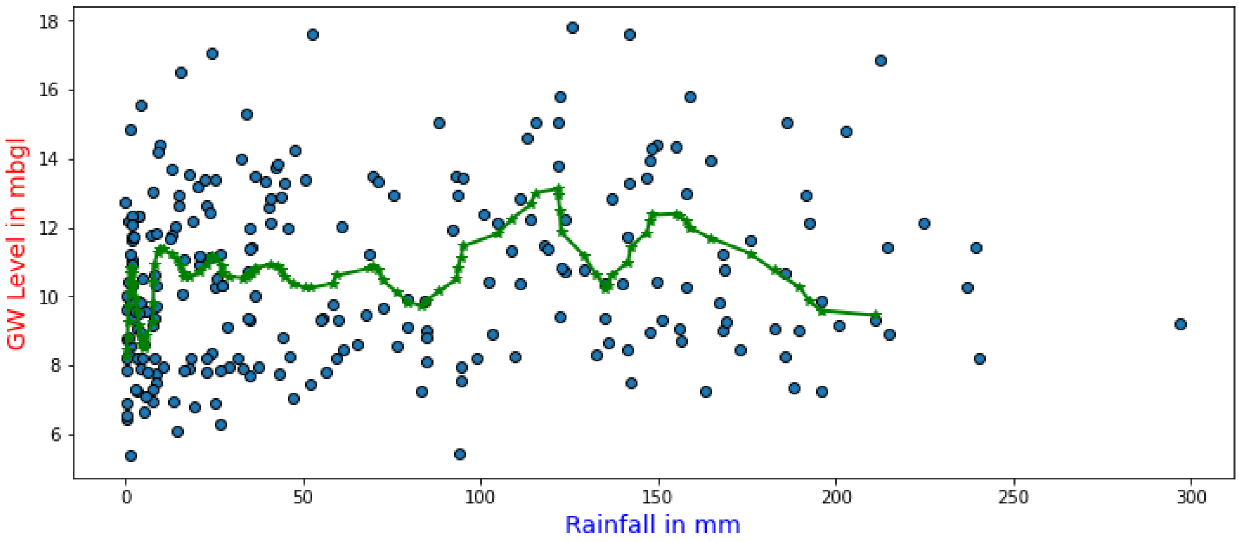

Fig. 6 Regression analysis of groundwater fluctuation with rainfall

R2 is the percentage of the dependent variable's variance that can be predicted by the independent variables [14]. The relationship between the rainfall and the ground water explored through the regression analysis is presented in Fig. 6 above. The regression analysis is performed using the gradient boosted regression algorithm. The R2 score estimated through the regression analysis between ground water and the rainfall in the study region is 0.9685. A positive association between groundwater level and rainfall was visible through the fitting curve shown in Fig. 6 .

CONCLUSION

Estimating groundwater recharge is essential in areas where groundwater is primarily used for water supply. In India, particularly in the state like Andhra Pradesh groundwater resources are heavily utilized for irrigation and other purposes and exposed to pollution. A regression model was created using Gradient Boosting Regression. With rainfall as inputs, the model was trained to estimate groundwater level. The factors causing groundwater variability in the studied area using the model can be determined. The actual values versus the estimated values of the groundwater level satisfied the linear relationship between the groundwater level and the primary components (R2 = 0.9685). Using a non-parametric method Mann-Kendal's test and theil slope estimator the similarity in the trends observed in the groundwater level and the rainfall in the state was analyzed and the variation in the trends helped to identify over exploited time period. It was observed that one of the main reasons for the groundwater level fluctuation in the study area was the intensive cultivation of water-intensive crops as well as variations in rainfall. Therefore, changing cropping patterns and artificially replenishing water supplies with effective irrigation techniques with proper management of water resources shall help to overcome impending hazardous conditions.

References

1. Yan, Shao-feng, et al. "Understanding groundwater table using a statistical model." Water Science and Engineering 11.1 (2018): 1-7.
2. Dangar, Swarup, Akarsh Asoka, and Vimal Mishra. "Causes and implications of groundwater depletion in India: A review." Journal of Hydrology 596 (2021): 126103.
3. Padmavathi, A. "GROWTH OF WASTE LAND IN THE STATEOF ANDHRA PRADESH." International journal of economic perspectives 16.5 (2022): 227-240.
4. Nageswara Rao, Kakani, et al. "Climate change and sea-level rise: Impact on agriculture along Andhra Pradesh coast—A geomatics analysis." Journal of the Indian Society of Remote Sensing 39.3 (2011): 415-422.
5. June 2022, Andhra Pradesh Water Resources Information and Management System. https://apwrims.ap.gov.in/
6. Büyükşahin, Ümit Çavuş, and Şeyda Ertekin. "Improving forecasting accuracy of time series data using a new ARIMA-ANN hybrid method and empirical mode decomposition." Neurocomputing 361 (2019): 151-163.

7. Alvarez-Ramirez, Jose, Eduardo Rodriguez, and Juan Carlos Echeverría. "Detrending fluctuation analysis based on moving average filtering." Physica A: statistical mechanics and its applications 354 (2005): 199-219.
8. Mondal, Arun, Sananda Kundu, and Anirban Mukhopadhyay. "Rainfall trend analysis by Mann-Kendall test: A case study of north-eastern part of Cuttack district, Orissa." International Journal of Geology, Earth and Environmental Sciences 2.1 (2012): 70-78.
9. Ahmed, Shamseddin Musa. "Assessment of irrigation system sustainability using the Theil–Sen estimator of slope of time series." Sustainability science 9.3 (2014): 293-302.
10. Friedman, Jerome H. "Stochastic gradient boosting." Computational statistics & data analysis 38.4 (2002): 367-378.
11. March 2022, Ministry of Jal Sakthi, http://jalshakti-dowr.gov.in/
12. Kumar, C. P., and P. V. Seethapathi. "Assessment of natural groundwater recharge in Upper Ganga Canal command area." Journal of Applied Hydrology 15.4 (2002): 13-20.
13. Shukla, Rituraj, et al. "Integrated water resources management in Sikta irrigation system, Nepal." Advances in Remediation Techniques for Polluted Soils and Groundwater. Elsevier, 2022. 337-359.
14. Colton, James A., and Keith M. Bower. "Some misconceptions about R2." International Society of Six Sigma Professionals, EXTRAOrdinary Sense 3.2 (2002): 20-22.