Fitting Of Gompertz Model Between Rainfall And Ground Water Levels – A Case Study

Raju Sake [1], P. Mohammed Akhtar[2]
[1] *Research Scholar, Department of Statistics, Sri Krishnadevaraya University, Anantapuramu,(A.P), India*
[2] *Professor, Department of Statistics, Sri Krishnadevaraya University, Anantapuramu,(A.P), India*

Abstract — *Time series analysis and its prediction itself involve tedious activities, such as their preprocessing, their transformation, and adjustment of various parameters and associated models etc. In this paper we have considered the Statistical Analysis of Rainfall and Ground Water Levels in Anantapuramu District of Andhra Pradesh. It deals with the application of Time Series model to analyze and predict Rainfall (RF) and Ground water levels (GWLs) in Anantapuramu district based on the data collected from January 2007 to December 2016. Through with Gompertz model for the purpose of analysis the district is divided into five zones. We have estimated the Gompertz model values and compared them by using the data. Further, validation of the fitted model identified the best suitable zone. i.e., least Mean Square Error (MSE) value of the zone and forecast on the Rainfall and Ground water levels of this district. We also find the relationship between rainfall and ground water levels in this district and conclusions are drawn based on the results obtained.*

Keywords — *Rainfall, Ground Water Level, Gompertz model, Validation of the model, prediction.*

I. INTRODUCTION

According to **Ya-lun Chou** "A time series may be defined as a collection of readings belonging to different time periods, of some economic variable or composite of variables." Mathematically, a time series is defined by the functional relationship

$$U_t = f(t) \qquad \text{……… (1.1)}$$

Where U_t is the value or the phenomenon (or variable) under consideration at time t.
For example, the temperature (U_t) of a place on different days (t) of the week; rainfall (U_t) of a place on different days (t) of the month; ground water levels (U_t) of a place on different months (t) of the year etc[3,4,5,6,11].

Named after **Benjamin Gompertz (1779-1865)** this is a sigmoid function. It is a useful model for time series in which growth is very slow in the initial and final stages. It is also useful in describing the growth of tumors, acquisition of mobile phones, and population in a confined space. Etc.

Gompertz Curve is used in actuarial work and sometimes in business and population forecasting as a growth curve. The function was originally designed to describe human mortality, but since it has been modified to be applied in biology, with regards to detailing populations [7, 8].

Water is the main source for any developmental activity. Water resources/water facilities are measured through Rainfall and Ground water levels in any region [9, 10,12]. Identifying the importance of Ground water level a time series analysis is proposed to analyze the data relating to Anantapur district because this district has been a drought prone area since so many decades [13]. In order to forecast Ground water levels of this district, at an attempt is made to predict Rainfall and Ground water levels through Modified exponential model [14].

In this direction, Average Ground Water Level (GWL) measured in meters (m) from 194 Piezometer points spread throughout the district and Average Rainfall measured in mille meters (m.m) of the district are considered. The data on the above variables are collected from the records of Ground Water and Water Audit Department Anantapuramu on Ground Water Levels (GWLs) and the data on Rainfall is collected from the Chief planning office, Anantapuramu from 2001 Jan to 2017 Oct. Further, Rainfall data is recorded on daily basis and Ground Water Levels are recorded on monthly basis from the respective records maintained by them [13, 14].

For the present work the data collected, the data relating to January to December months from 2007 to 2016 is considered for the purpose of analysis of this paper on both the variables i.e. Ground Water Level and Rainfall. Further Anantapuramu district consisting of 63 mandals is divided into **five** Revenue Divisions for the administrative convenience and hence for the analysis these five Revenue Divisions are considered as five zones and are given in the following table along with their respective Mandals [1,2,13,14].

Table-1.1

Zonal-wise (Revenue division) of mandals in Anantapuramu District.

Sl.No	Zone-I Anantapuramu RD	Zone-II Penukonda RD	Zone-III Kadiri RD	Zone-IV Kalyandurg RD	Zone-V Dharmavaram RD
1	Anantapuramu	Agali	Amadagur	Beluguppa	Bathalapalli
2	Atmakur	Amarapuram	Bukkapatnam	Bommanahal	C.K.Palli
3	B.K.Samudram	Chilamathur	Gandlapenta	Brahmasamudram	Dharmavaram
4	Garladinne	Gorantla	Kadiri	D.Hirehal	Kanaganipalli
5	Gooty	Gudibanda	Kothacheruvu	Gummaghatta	Mudigubba
6	Guntakal	Hindupur	N.P.Kunta	Kalyandurg	Ramagiri
7	Kudair	Lepakshi	Nallacheruvu	Kambadur	Raptadu
8	Narpala	Madakasira	Nallamada	Kanekal	Tadimarri
9	Pamidi	Parigi	O.D.Cheruvu	Kundurpi	
10	Peddapappur	Penukonda	Puttaparthi	Rayadurg	
11	Peddavadugur	Roddam	Talupula	Settur	
12	Putlur	Rolla	Tanakal		
13	Singanamala	Somandepalli			
14	Tadipatri				
15	Uravakonda				
16	Vajrakarur				
17	Vidapanakal				
18	Yadiki				
19	Yellanur				
Total(63)	19	13	12	11	8

Similarly, zonal wise Piezometer Points are also provided in the following table, from which GWLs are measured.

Table-1.2
Zonal-wise of Piezometer Points in Anantapuramu District.

	Zone-I Anantapuramu RD	Zone-II Penukonda RD	Zone-III Kadiri RD	Zone-IV Kalyandurg RD	Zone-V Dharmavaram RD
Piezometer Points(194)	54	50	31	32	27

The data is collected on Average Rainfall and Average Ground Water Levels are given in the following Table-1.3 for a ready reference.

Table-1.3
Average Rainfall and Average Ground water levels data from 2007 to 2016

Year	**Zone-I**		**Zone-II**		**Zone-III**		**Zone-IV**		**Zone-V**	
	RF (in mm)	GWL	RF (in mm)	GWL	RF (in mm)	GWL	RF (in mm)	GWL	RF (in mm)	GWL
2007	65.60	10.57	58.20	22.58	67.20	14.23	52.00	14.97	60.50	17.03
2008	53.90	9.96	77.90	20.73	65.20	9.27	61.30	10.88	62.70	9.09

2009	45.40	12.17	50.60	17.53	46.30	11.08	57.10	9.58	38.70	10.24
2010	53.90	12.74	71.50	15.02	70.80	12.03	64.60	8.58	56.30	11.79
2011	39.50	12.69	42.30	15.20	48.90	11.48	31.80	8.93	36.60	12.84
2012	43.20	14.98	43.40	20.49	45.30	16.08	40.50	13.76	41.90	13.22
2013	35.00	15.94	52.30	23.03	47.10	18.69	34.80	16.98	38.10	14.30
2014	31.10	15.87	30.30	23.40	27.10	21.16	37.10	18.92	22.80	16.30
2015	44.10	14.90	62.60	26.88	66.30	25.80	46.00	19.26	54.30	17.66
2016	33.50	15.57	33.40	27.27	32.30	15.35	25.70	19.51	30.10	16.15

II. STATISTICAL ANALYSIS

Some of the Preliminary Statistical analysis is done for the data provided in the above table - 1.3, such as yearly averages of Rainfall and Ground water levels are calculated and Karl-Pearson's Correlation Coefficient (r) is calculated between Average Rainfall(X) and Average Ground water levels (Y) Zonal wise by using the following formula,

$$\mathbf{r} = \frac{cov(x,y)}{\sqrt{v(x).v(y)}} \quad ....(2.1)$$

and are given in the following Table-2.1.

Table-2.1
Correlation Coefficient between Average Rainfall and Average Ground Water Level.

Years	Zone-I	Zone-II	Zone-III	Zone-IV	Zone-V
2007-2016	-0.84	-0.26	-0.20	-0.58	-0.23

By studying the above Correlation Coefficients we can observed that all the Correlation Coefficients are negative, that is the relation between Rainfall and Ground Water levels is negative, that is, if Rainfall is increases the Ground water level decreases, it is true, because the depth of the water level will decrease. By observing the Correlation coefficients in the above Table-2.1 in Zone-I and Zone-IV they are strongly negatively related, as the other Zones are weakly negatively related. We observed that in Zone-I and Zone-IV additional to rainfall, other water resources like, High Level Canal (HLC) in these zones that also helps to improve the Ground water level.

To forecast **Rainfall** and **Ground Water Levels** through Gompertz model for different zones, we consider

The Gompertz Model $\boldsymbol{y_t = k * a^{b^t}}$ ....(2.2)

$$\log_e y_t = \log_e k + b^t \log_e a \quad ....(2.3)$$

$$Y_t = K + Ab^t \quad ....(2.4)$$

$$a = antilog(A), k = antilog(K) \quad ....(2.5)$$

Where $b = (\frac{y_3-y_2}{y_2-y_1})^{(\frac{1}{t_2-t_1})}$ ....(2.6)

$$A = \log_e a = a = \frac{(y_2-y_1)^2}{(y_3-2y_2+y_1)} * (\frac{y_2-y_1}{y_3-y_2})^{(\frac{t_1}{t_2-t_1})} \quad ......(2.7)$$

$$k = \log_e k = \frac{(y_1y_3-y_2^2)}{(y_3-2y_2+y_1)} \quad ......(2.8)$$

Here, t_1, t_2, t_3 are three selected time points and
y_1, y_2, y_3 are their correspondence rainfall or ground water level values it's taken.

To fit the above Gompertz model and to estimate the values of the parameters 'a', 'b' and 'k' by solving the related normal equations and following trend curve is fitted for the data given in table 1.3 and fitted model is given below.

The fitted Gompertz model for Average RF and Average GWLs:

A: For Rainfall

Zone-I
Gompertz Curve $\log_e y_t = (2.62) + (1.63) * (0.92)^t$

Zone-II
Gompertz Curve $\log_e y_t = (3.03) + (2.02) * (0.81)^t$

Zone-III
Gompertz Curve $\log_e y_t = (4.46) + (-0.17) * (1.27)^t$

Zone-IV
Gompertz Curve $\log_e y_t = (3.58) + (1.44) * (-0.61)^t$

Zone-V
Gompertz Curve $\log_e y_t = (-0.03) + (4.57) * (0.95)^t$

B: For Ground water levels

Zone-I
Gompertz Curve $\log_e y_t = (5.18) + (-3.05) * (0.97)^t$

Zone-II
Gompertz Curve $\log_e y_t = (2.90) + (0.10) * (-1.12)^t$

Zone-III
Gompertz Curve $\log_e y_t = (2.12) + (0.05) * (1.43)^t$

Zone-IV
Gompertz Curve $\log_e y_t = (2.35) + (0.02) * (-1.55)^t$

Zone-V
Gompertz Curve $\log_e y_t = (3.37) + (-1.46) * (0.89)^t$

Gompertz Curve $\hat{y}_t = \hat{e}(\log_e y_t)$ here substitutes the $\boldsymbol{log_e y_t}$ values for required estimated Gompertz curve values.

III. VALIDATION OF THE FITTED MODEL

Validation of the fitted model is necessary to check the suitability of the model for the given data and which is done by considering X = Years and Y = Average RF or Average GWL given in table-1.3 and estimated the Average RF (Y) or Average GWL (Y) denoted by $\hat{y}$. The estimated Average RF and Average GWL are given in the following tables.

Table-3.1

Estimated Average RF $\hat{y}$ for Gompertz Curve.

Year	Zone-I		Zone-II		Zone-III		Zone-IV		Zone-V	
	Actual	Estimates	Actual	Estimates	Actual	Estimates	Actual	Estimates	Actual	Estimates
2007	65.60	61.56	58.20	106.70	67.20	69.41	52.00	14.88	60.50	74.44
2008	53.90	55.15	77.90	78.26	65.20	66.02	61.30	60.95	62.70	59.15
2009	45.40	48.91	50.60	60.34	46.30	60.95	57.10	25.79	38.70	49.40
2010	53.90	44.26	71.50	49.40	70.80	55.70	64.60	43.82	56.30	39.25
2011	39.50	40.45	42.30	42.10	48.90	49.40	31.80	31.82	36.60	32.79
2012	43.20	36.97	43.40	36.60	45.30	42.52	40.50	38.47	41.90	28.50
2013	35.00	34.12	52.30	32.79	47.10	34.81	34.80	34.47	38.10	23.81
2014	31.10	31.50	30.30	30.27	27.10	27.39	37.10	36.97	22.80	19.89
2015	44.10	29.67	62.60	27.94	66.30	20.09	46.00	35.52	54.30	17.29
2016	33.50	27.66	33.40	26.31	32.30	13.46	25.70	36.23	30.10	15.03

Table-3.2

Estimated Average GWL $\hat{y}$ for Gompertz Curve.

Year	Zone-I		Zone-II		Zone-III		Zone-IV		Zone-V	
	Actual	Estimates	Actual	Estimates	Actual	Estimates	Actual	Estimates	Actual	Estimates
2007	10.57	9.21	22.58	16.28	14.23	8.94	14.97	10.18	17.03	7.92
2008	9.96	10.07	20.73	20.70	9.27	9.21	10.88	11.02	9.09	9.21
2009	12.17	11.02	17.53	15.80	11.08	9.68	9.58	9.78	10.24	10.49
2010	12.74	11.82	15.02	21.33	12.03	10.28	8.58	11.82	11.79	11.59
2011	12.69	12.94	15.20	15.18	11.48	11.25	8.93	8.76	12.84	12.81
2012	14.98	14.15	20.49	22.20	16.08	12.81	13.76	13.87	13.22	14.01
2013	15.94	15.03	23.03	14.59	18.69	15.33	16.98	6.82	14.30	15.33
2014	15.87	16.44	23.40	23.34	21.16	19.89	18.92	20.49	16.30	16.44
2015	14.90	17.46	26.88	13.74	25.80	29.08	19.26	3.74	17.66	17.46
2016	15.57	18.54	27.27	24.78	15.35	49.90	19.51	51.94	16.15	18.54

In the above tables-3.1 and 3.2 for the validation of the model Mean Square Errors (MSE's) are calculated zone wise by considering

$$\textbf{MSE} = \sum(y-\hat{y})^2 \quad \text{.....(3.1)}$$

Where y represents actual or observed values given in table-1.3 and $\hat{y}$ is the estimated values through fitted Gompertz model is given in tables-3.1 and 3.2 using fitted Gompertz model respectively. MSE's were calculated and are given in the following Table-3.3.

Table-3.3

MSE's for Average RF- Gompertz Model.

Type of the Model	Zone-I	Zone-II	Zone-III	Zone-IV	Zone-V
Gompertz	406.11	4614.16	3097.61	3015.10	2615.71

Table-3.4

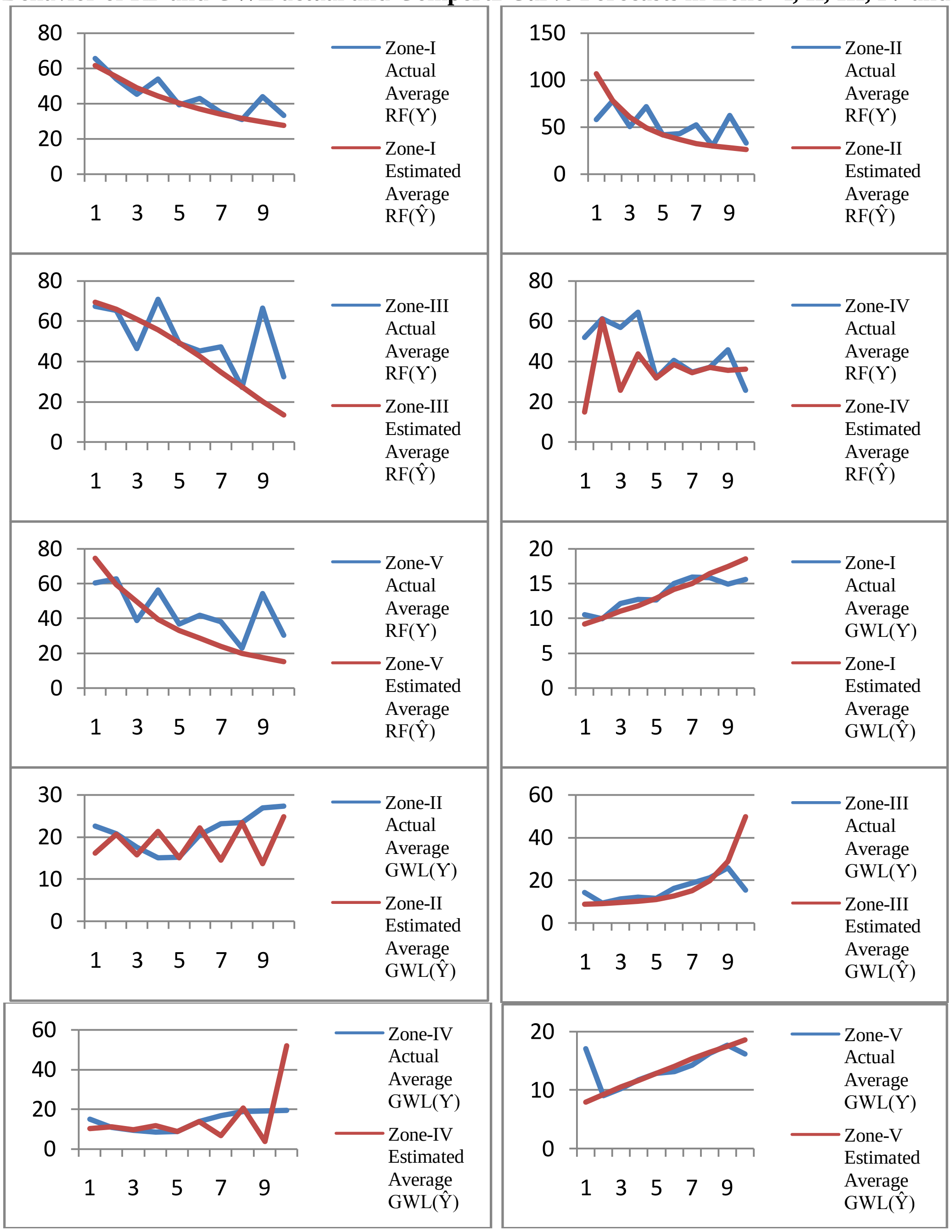
MSE's for Average GWL – Gompertz Model.

Type of the Model	Zone-I	Zone-II	Zone-III	Zone-IV	Zone-V
Gompertz	21.31	335.52	1261.12	1431.81	90.57

By Comparing MSE's for RF and GWLs through Gompertz model under consideration, for RF of zone-I is least and GWLs for zone-I Gompertz model is the most suitable model because MSEs for zone-I is least. Next to zone-I, zone-V has least MSEs. Thus next to zone-I for zone-V Gompertz model is best suitable for the RF and GWLs. Further, the behaviors of RF and GWL through this model i.e. Gompertz model in different Zones are represented in the following Figure-3.1. Similar conclusions can be drawn from the following graphs also.

Fig-3.1

Behavior of RF and GWL actual and Gompertz Curve Forecasts in Zone –I, II, III, IV and V



Note: In the above graphs x-axis represents years in the last decade i.e. from 2007 to 2016.
On y-axis RF measured in Mille Meters or Average GWLs measured in Meters.

IV. FURTHER STATISTICAL ANALYSIS

Now we proceed to analyze the given estimates in tables-3.1 and 3.2 using ANOVA two- way classification by considering rows as different years and columns as different zones and the following Null Hypothesis are formed and tested.

H_{01}: There is no significant difference between different years of Average RF in Anantapuramu District.
H_{02} : There is no significant difference between Average RF of different zones in Anantapuramu District.
H_{03}: There is no significant difference between different years of Average Ground Water Levels in Anantapuramu District.
H_{04} : There is no significant difference between Average Ground Water Levels of different zones in Anantapuramu District.

Table-4.1
ANOVA Two-way Table for RF

Source of variation	d.f	S.S	M.S.S	F-cal
Rows (years)	9	9885.685	1098.409	7.13**
Columns (Zones)	4	1253.674	313.4186	2.03*
Error	36	5538.276	153.841	
Total	49	16677.64		

By comparing F-calculated value of Rows (Years) with F-critical value at 5 % level of significance (l.o.s) we reject the H_{01} i.e. There is a significant difference between different years of Average RF in Anantapuramu District. Similarly by comparing F-calculated value of Columns (Zones) with F-critical value at 5 % level of significance (l.o.s) we accept the H_{02} i.e. There is no significant difference between different zones of Average RF in Anantapuramu District.

Table-4.2
ANOVA Two-way Table for GWL

Source of variation	d.f	S.S	M.S.S	F-cal
Rows (years)	9	1930.554	214.506	4.70*
Columns (Zones)	4	235.6598	58.91496	1.29*
Error	36	1641.98	45.61056	
Total	49	3808.194		

By comparing F-calculated value of Rows (Years) with F-critical value at 5 % level of significance (l.o.s) we reject the H_{01} i.e. There is a significant difference between different years of Average GWL in Anantapuramu District. Similarly by comparing F-calculated value of Columns (Zones) with F-critical value at 5 % level of significance (l.o.s) we accept the H_{02} i.e. There is no significant difference between different zones of Average GWL in Anantapuramu District.

Since F-cal value related to rows(years) in RF and rows(years) in GWL is high so there is a necessity for Critical Difference (C.D) Test for sub-grouping various years using the following formula[11][12].

C.D. = $\sqrt{2 \times Error\ M.S.S/m}$ × $t_{0.01}$ for error d.f. in tables -4.1 and 4.2 ......(4.1)

Where m represents number of estimates in each zone and as well as year.

V. CRITICAL DIFFERENCE (C.D) TEST: Average RF for Years

Table-5.1

Year wise Aggregate Average RF for Gompertz Curve estimates

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Average	65.39	63.90	49.07	46.48	39.31	36.61	32	29.20	26.10	23.73
Ranking	X	IX	VIII	VII	VI	V	IV	III	II	I

Table 5.2
If we can arranged Ascending Order

Year	2016	2015	2014	2013	2012	2011	2010	2009	2008	2007
Average	23.73	26.10	29.20	32	36.61	39.31	46.48	49.07	63.90	65.39

S.E = $\sqrt{2 \times Error\ M.S.S/m}$
= 7.84

1% l.o.f C.D = 2.58×7.84
= 20.22

2016 2015 2014 2013 2012 2011 2010 2009 2008 2007

Above notation indicates that 2016, 2015, 2014, 2013, 2012 ,2011 years Average RF come under one category and 2015, 2014, 2013, 2012, 2011 years Average RF and 2014, 2013, 2012, 2011, 2010, 2009 years Average RF and also 2010, 2009, 2008, 2007 come under another category because there is no Significant Difference in average RF. These years are ranked based on their respective Average RF.

CRITICAL DIFFERENCE (C.D) TEST: Average GWL for Years

Table-5.3
Year wise Aggregate Average Ground Water Levels for Gompertz Curve estimates

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Average	10.50	12.04	11.35	13.36	12.18	15.40	13.42	19.32	16.29	32.74
Ranking	I	III	II	V	IV	VII	VI	IX	VIII	X

Table 5.4
If we can arranged Ascending Order

Year	2007	2009	2008	2011	2010	2013	2012	2015	2014	2016
Average	10.50	11.35	12.04	12.18	13.36	13.42	15.40	16.29	19.32	32.74

S.E = $\sqrt{2 \times Error\ M.S.S/m}$
= 4.27

1% l.o.f C.D = 2.58×4.27
= 11.01

2007 2009 2008 2011 2010 2013 2012 2015 2014 2016

Above notation indicates that 2007, 2009, 2008, 2011, 2010, 2013, 2012, 2015, 2014 Average GWLs come under one category and 2016 Average GWLs, come under another category because there is no Significant Difference in average ground water levels. These years are ranked based on their respective average GWLs.

ACKNOWLEDGEMENTS

The authors express heartfelt thanks to the persons responsible for giving the necessary data on Rainfall and Ground Water Levels working in Ground Water and Water Audit Department and Chief Planning Office Anantapuramu. Further the authors are profusely thankful to Mr. C. Pothulaiah, Asst. Hydro geologist, Ground Water and Water Audit Department and Mr. Muralimohan Reddy, Assitant Statistical Officer, Chief Planning Office, Anantapuramu and also Mr. N.V. Subbarao, Mr. Ramakrishna and finally Mr. Muralikrishna from Jaipur in Rajasthan (State) for his timely suggestions and useful discussions.

AUTHORS PROFILE

Mr. Raju Sake is doing research in Statistics in the Department of Statistics, S.K. University, Anantapur, for his Ph.D.

Prof. P.Mohammed Akhtar, working as Professor of Statistics, in the Department of Statistics, S.K. University, Anantapur, for the last 33 years. He has published more than 40 research papers in various journals. He produced 6 Ph.D's under his guidance. His main research work focuses on Quality Control, Time Series Analysis, and Data Mining.

REFERENCES

[1] A.H. Nury, M. Koch and M.J.B. Alam, "*Time Series Analysis and Forecasting of Temperatures in the Sylhet Division of Bangladesh*", Proceedings of 4th International Conference on Environmental Aspects of Bangladesh, Fukoka, Japan, 2013.

[2] D. Machiwal and M.K. Jha, "*Time Series Analysis of Hydrologic Data For Water Resources Planning and Management*": A Review. J. Hydrol, Hydromech, 54(3): 237-257, 2006.

[3] D.C Montgomery, L.A Johnson," *Forecasting and time series analysis"*. McGraw-Hill; 1976.

[4] G.E.P Box, G.M. & G.C Reinsel, "*Time series analysis and forecasting and control"*, 3rd ed., Englewood Cliffs, N.J. Prentice Hall, 1994.

[5] G.E.P Box, G.M Jenkins, *"Time Series Analysis, Forecasting and Control"*, Holden-Day: San Francisco, 1976.

[6] G.E.P Box, G.M Jenkins, G.C Reinsel, *"Time series analysis: Forecasting and Control"*. John Wiley and Sons; 2008.

[7] J.D Hamilton, *"Time series analysis"*. Princeton University Press; 1994.

[8] Otnes RK, Enochson L. *"Applied Time Series Analsis"*, Vol. 1. New York: Wiley; 1978.

[9] S. Soltani, R. Modarres and S.S. Eslamian, " *The use of time series modeling for the determination of rainfall climates of Iran*". International Journal of Climatology, 27: 819-829, 2007.

[10] S.A. Shamsnia, N. Shahidi, A. Liaghat, A. Sarraf and S.F. Vahdat, " *Modelling Of Weather Parameters (Temperature, Rainfall And Humidity) Using Stochastic Methods*. Internat. Conference on Environment and Industrial Innovation, IPCBEE, Singapore, 282-285, 2011.

[11] S.C. Gupta, V.K. Kapoor, *"Fundamentals of Applied Statistics".* Sultan Chand & Sons; 2003.

[12] S.M. Ali, "Time *Series Analysis of Baghdad Rainfall Using ARIMA method* "Iraqi Journal of Science,54(4):1136-1142,2013.

[13] S. Raju, P. Mohammed Akhtar, *"Time Series Analysis on Rainfall and Ground Water Levels Data – A Case Study",* International Journal of Scientific Research in Mathematics and Statistical Sciences Vol.6, Issue.1, pp.76-85, February (2019).

[14] S. Raju, P. Mohammed Akhtar, *"Fitting of modified exponential model between rainfall and ground water levels: A case study".* International Journal of Statistics and Applied Mathematics 2019; 4(4): 01-06.