International Journal of Engineering & Scientific Research
Vol.9, Issue 6, June 2021,
ISSN: 2347-6532 Impact Factor: 6.660
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Fitting of Poisson distribution between Rainfall and Ground water levels – A Case Study

Raju Sake
Academic Consultant, Department of Statistics, Sri Krishnadevaraya University, Anantapuramu (A.P), India
E-mail: raju.hcu2007@gmail.com, *Tel.: 9490715718*

Abstract—Present paper deals with the application of Distribution Theory 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 Poisson distribution for the purpose of analysis the district is divided into five zones or Revenue Divisions (RD) namely, 1. Anantapuramu RD 2. Penukonda RD 3. Kadiri RD 4. Kalyandurg RD 5. Dharmavaram RD. I have estimated the Poisson distribution values and compared among them by using the data. Further, validation of the fitted distribution identified the best suitable zone. i.e., Residual Analysis or Error Analysis or Residual Sum of Squares (RSS) or Error Sum of Squares (ESS) or least Mean Square Error (MSE) value of the zone and forecast on the Rainfall and Ground water levels of this district. We also calculate Critical Difference (C.D) test and conclusions are drawn based on the results obtained.

Keywords—*Rainfall, Ground Water Level, Residual Analysis, Validation of the distribution, C.D test.*

1. INTRODUCTION

Earlier we have discussed in the previous paper [1] the method of curve fitting is the best for estimating trend. The nature of the curve that is appropriate for the given data can be satisfactorily decided either by a theoretical understanding of the data or by observing the scatter diagram that is constructed for the given data.

The methods of fitting Straight Line, Second Degree Parabola, Exponential Curve and Power Curves by least squares method was discussed [1].

Linear, Parabolic, Exponential and Power Curve projections generally assume that growth or decline continues without limit. While these trends continue for some time they are not continue forever. There are a number of situations in which there is an asymptote to growth or decline. There are three types of Growth Curves or Models is there, that is:

1. Modified Exponential Model [2, 6]
2. Gompertz Model [3, 5]
3. Logistic Model. These models also discussed earlier papers in 'Time Series Analysis and Forecasting' concept [4].

I have discussed 'Distribution Theory' for different distributions like. Binomial Distribution-Direct and Recurrence Relation Method and Negative Binomial Distribution Recurrence Relation Method already I have analyzed, now I will fit **Poisson distribution** in this paper.

International Journal of Engineering & Scientific Research
Vol.9, Issue 6, June 2021,
ISSN: 2347-6532 Impact Factor: 6.660
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

The data is collected on Average Rainfall and Average Ground Water Levels are given in the following Table-1.1 for a ready reference [1, 2, 3, 4, 5, 6, 7 and 8].

Table-1.1
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

2. STATISTICAL ANALYSIS

Some of the Preliminary Statistical analysis is done for the data provided in the above table -1.1, 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[1, 2 and 3].

To forecast **Rainfall** and **Ground Water Levels** through Poisson distribution for different zones we can consider given as follows:

The Probability Mass Function of Poisson distribution is given by

$$P(X = x) = \frac{e^{-\lambda}\lambda^{x}}{x!} \quad where\ x = 1,2,3, \ldots, \qquad \ldots\ldots (2.1)$$

In Poisson distribution the parameter λ is equal to the Arithmetic mean.

$$\bar{x} = \frac{\sum_{i=1}^{n} f_i\, x_i}{N} = \lambda = \hat{\lambda}$$

Substitute this $\hat{\lambda}$ in the above Poisson distribution equation (2.1) we will get the Probability Mass Function of a Poisson distribution has

$$P(X = x) = \frac{e^{-\hat{\lambda}}\hat{\lambda}^{x}}{x!} \quad where\ x = 1,2,3, \ldots, \qquad \ldots\ldots. (2.2)$$

International Journal of Engineering & Scientific Research
Vol.9, Issue 6, June 2021,
ISSN: 2347-6532 Impact Factor: 6.660
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

To find the Expected Frequencies: To find the Expected Frequencies, we use the following probability Mass function of Poisson distribution is given by

$P(X = x) = \frac{e^{-\hat{\lambda}}\hat{\lambda}^x}{x!}$....... (2.3)

The fitted Poisson distribution for Average RF and Average GWLs:

A: For Average Rainfall

Zone-I
The Probability Mass Function of Poisson distribution is given by

$$P(X = x) = \frac{e^{-4.96}(4.96)^x}{x!}$$

Zone-II
The Probability Mass Function of Poisson distribution is given by

$$P(X = x) = \frac{e^{-5.03}(5.03)^x}{x!}$$

Zone-III
The Probability Mass Function of Poisson distribution is given by

$$P(X = x) = \frac{e^{-5.04}(5.04)^x}{x!}$$

Zone-IV
The Probability Mass Function of Poisson distribution is given by

$$P(X = x) = \frac{e^{-4.92}(4.92)^x}{x!}$$

Zone-V
The Probability Mass Function of Poisson distribution is given by

$$P(X = x) = \frac{e^{-4.98}(4.98)^x}{x!}$$

B: For Average Ground water levels

Zone-I
The Probability Mass Function of Poisson distribution is given by

$$P(X = x) = \frac{e^{-5.91}(5.91)^x}{x!}$$

Zone-II

The Probability Mass Function of Poisson distribution is given by

$$P(X = x) = \frac{e^{-5.84}(5.84)^x}{x!}$$

Zone-III

The Probability Mass Function of Poisson distribution is given by

$$P(X = x) = \frac{e^{-6.15}(6.15)^x}{x!}$$

Zone-IV

The Probability Mass Function of Poisson distribution is given by

$$P(X = x) = \frac{e^{-6.12}(6.12)^x}{x!}$$

Zone-V

The Probability Mass Function of Poisson distribution is given by

$$P(X = x) = \frac{e^{-5.83}(5.83)^x}{x!}$$

$where\ x = 1,2,3 \ldots$ Substitute in the above equations we can get the values of $p(1), p(2), p(3), \ldots, multiplying\ these\ p(1), p(2), p(3), \ldots,\ values\ by\ the\ N = \sum_{i=1}^{n} f_i$ we get the required Expected Frequencies, these are denoted by $f(1), f(2), f(3), \ldots,$

3. VALIDATION OF THE FITTED DISTRIBUTION

Validation of the fitted distribution is necessary to check the suitability of the distribution for the given data this is done by considering X = Years and Y = Average RF or Average GWL given in table-1.1 and estimated the Average RF (Y) or Average GWL (Y) denoted by $\hat{y}$. The estimated Average RF and Average GWLs are given in the following tables.

International Journal of Engineering & Scientific Research
Vol.9, Issue 6, June 2021,
ISSN: 2347-6532 Impact Factor: 6.660
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

Table-3.1
Estimated Average RF $\hat{y}$ for Poisson distribution

Year	Zone-I		Zone-II		Zone-III		Zone-IV		Zone-V	
	Actual	Estimates	Actual	Estimates	Actual	Estimates	Actual	Estimates	Actual	Estimates
2007	65.60	13.36	58.20	15.68	67.20	15.50	52.00	18.04	60.50	13.26
2008	53.90	40.07	77.90	41.80	65.20	41.32	61.30	40.58	62.70	39.78
2009	45.40	62.33	50.60	73.15	46.30	72.31	57.10	63.13	38.70	61.88
2010	53.90	80.14	71.50	88.83	70.80	87.81	64.60	81.16	56.30	79.56
2011	39.50	80.14	42.30	94.05	48.90	92.97	31.80	81.16	36.60	79.56
2012	43.20	66.78	43.40	78.38	45.30	77.48	40.50	63.13	41.90	66.30
2013	35.00	44.52	52.30	57.48	47.10	56.82	34.80	45.09	38.10	44.20
2014	31.10	26.71	30.30	36.58	27.10	36.16	37.10	27.05	22.80	26.52
2015	44.10	17.81	62.60	20.90	66.30	20.66	46.00	13.53	54.30	17.68
2016	33.50	8.90	33.40	10.45	32.30	10.33	25.70	9.02	30.10	8.84

Table-3.2
Estimated Average GWL $\hat{y}$ for Poisson distribution

Year	Zone-I		Zone-II		Zone-III		Zone-IV		Zone-V	
	Actual	Estimates	Actual	Estimates	Actual	Estimates	Actual	Estimates	Actual	Estimates
2007	10.57	2.71	22.58	4.24	14.23	1.55	14.97	1.41	17.03	2.77
2008	9.96	6.77	20.73	10.61	9.27	6.21	10.88	5.65	9.09	6.93
2009	12.17	12.19	17.53	21.21	11.08	12.41	9.58	11.31	10.24	13.86
2010	12.74	18.95	15.02	29.70	12.03	20.17	8.58	18.38	11.79	19.41
2011	12.69	21.66	15.20	33.94	11.48	24.83	8.93	22.62	12.84	22.18
2012	14.98	21.66	20.49	33.94	16.08	24.83	13.76	22.62	13.22	22.18
2013	15.94	18.95	23.03	27.58	18.69	21.72	16.98	19.79	14.30	18.02
2014	15.87	13.54	23.40	21.21	21.16	17.07	18.92	15.55	16.30	13.86
2015	14.90	9.48	26.88	12.73	25.80	10.86	19.26	9.90	17.66	8.32
2016	15.57	5.42	27.27	8.49	15.35	7.76	19.51	5.65	16.15	5.54

In the above tables -3.1 and 3.2 for the validation of the distribution, Residual Analysis or Error Analysis or Residual Sum of Squares (RSS) or Error Sum of Squares (ESS) or Mean Square Errors (MSE's) are calculated zone wise by considering

$$Residual\ Sum\ of\ Squares\ (RSS) = \sum_{i=1}^{n} (y_i - \hat{y}_i)^2 \qquad \text{..... (3.1)}$$

Where $y_i\ or\ o_i$ represents actual or observed values given in table-1.1 and $\hat{y}_i\ or\ \hat{e}_i$ is the estimated values through fitted distribution is given in tables- 3.1 and 3.2. Residual Sum of Squares was calculated and is given in the following table.

Table-3.3
RSS values for Average RF for Poisson distribution.

Type of the Distribution	Zone-I	Zone-II	Zone-III	Zone-IV	Zone-V
Poisson distribution	7509.30	10153.51	9928.96	6381.12	8120.27

International Journal of Engineering & Scientific Research
Vol.9, Issue 6, June 2021,
ISSN: 2347-6532 Impact Factor: 6.660
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

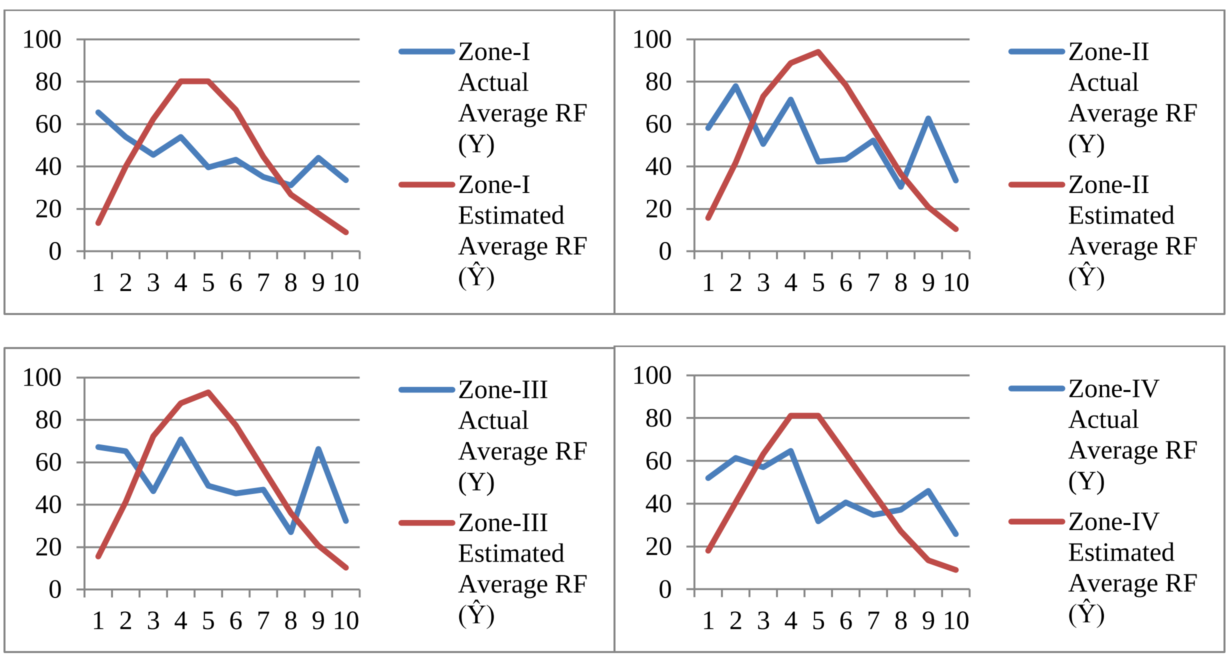
Table-3.4
RSS values for Average GWLs for Poisson distribution.

Type of the Distribution	Zone-I	Zone-II	Zone-III	Zone-IV	Zone-V
Poisson distribution	382.49	1778.31	799.67	875.14	666.30

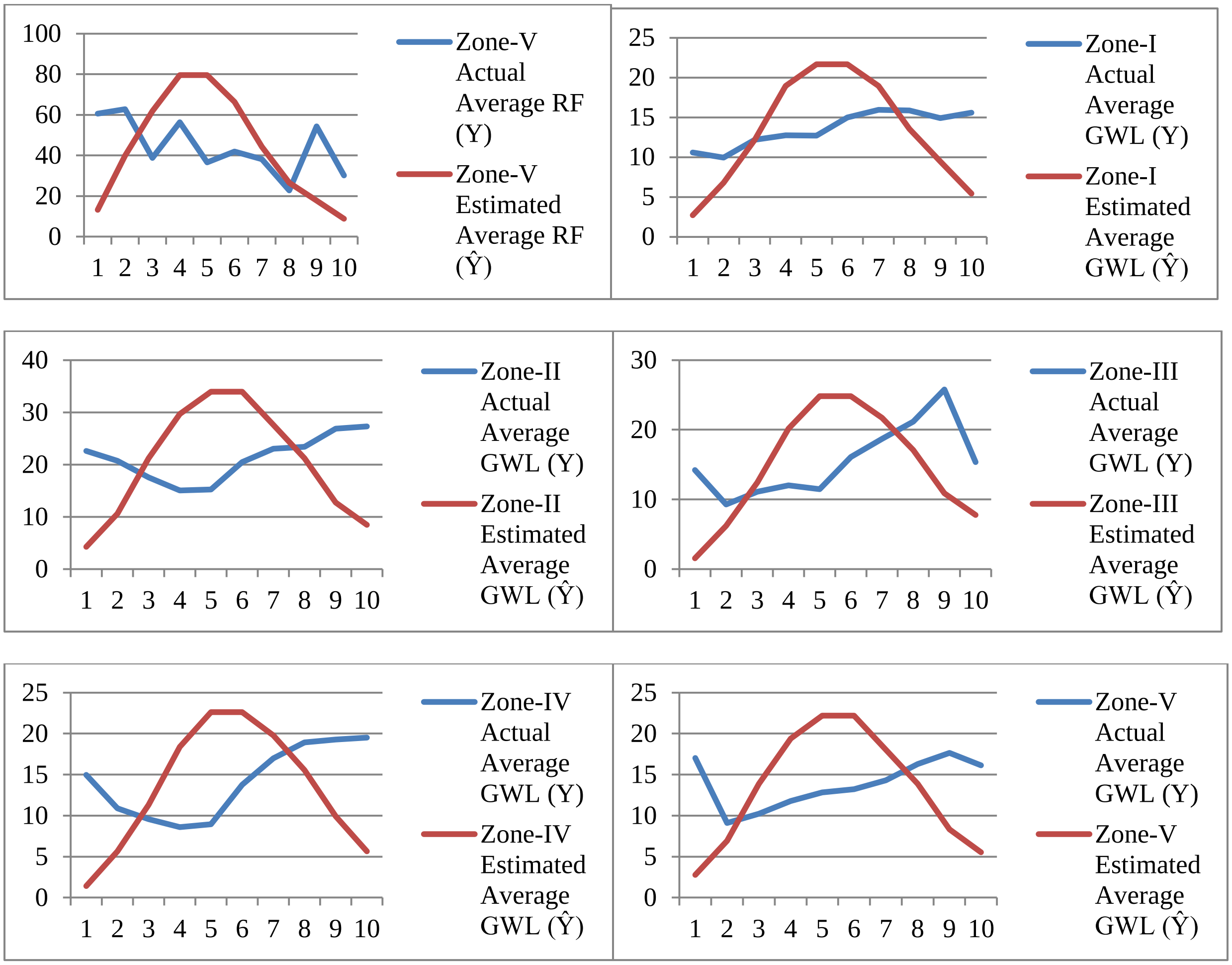
CONCLUSIONS:

By Comparing Residual Sum of Squares values for Average RF and Average GWLs through Poisson distribution under consideration, for RF of Zone-IV is least and GWLs for Zone-I Residual Sum of Squares values is least. Next to Zone-IV, Zone-I have least Residual Sum of Square value in RF and GWLs Zone-V is least. Further, the behaviors of RF and GWL through this distribution 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 GWLs Actual and Estimated values for Poisson distribution in Zone –I, II, III, IV and V



International Journal of Engineering & Scientific Research
Vol.9, Issue 6, June 2021,
ISSN: 2347-6532 Impact Factor: 6.660
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A



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

4. 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 [1, 2, 3, 4, 5, 6, 7 and 8].

H_{02}: There is no significant difference between Average RF of different zones in Anantapuramu District [1, 2, 3, 4, 5, 6, 7 and 8].

H_{03}: There is no significant difference between different years of Average Ground Water Levels in Anantapuramu District [1, 2, 3, 4, 5, 6, 7 and 8].

H_{04}: There is no significant difference between Average Ground Water Levels of different zones in Anantapuramu District [1, 2, 3, 4, 5, 6, 7 and 8].

Table-4.1
ANOVA Two-way Table for RF

Source of variation	d.f	S.S	M.S.S	F-cal
Rows (years)	9	36598.21	4066.467	474.7628
Columns (Zones)	4	664.3899	166.0975	19.39199
Error	36	308.3494	8.565261	
Total	49	37570.95		

By comparing F-calculated value of Rows (Years) with F-critical value at 5% level of significance 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 we reject the H_{02} i.e. There is a significant difference between Average RF of different zones in Anantapuramu District.

Table-4.2
ANOVA Two-way Table for GWLs

Source of variation	d.f	S.S	M.S.S	F-cal
Rows (years)	9	2958.843	328.7603	99.2878
Columns (Zones)	4	381.2768	95.31919	28.78703
Error	36	119.2027	3.311185	
Total	49	3459.322		

By comparing F-calculated value of Rows (Years) with F-critical value at 5% level of significance we reject the H_{03} i.e. There is a significant difference between different years of Average GWLs in Anantapuramu District. Similarly by comparing F-calculated value of Columns (Zones) with F-critical value at 5% level of significance we reject the H_{04} i.e. There is a significant difference between Average GWLs of different zones in Anantapuramu District.

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

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

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

International Journal of Engineering & Scientific Research
Vol.9, Issue 6, June 2021,
ISSN: 2347-6532 Impact Factor: 6.660
Journal Homepage: http://www.ijmra.us, Email: editorijmie@gmail.com
Double-Blind Peer Reviewed Refereed Open Access International Journal - Included in the International Serial Directories Indexed & Listed at: Ulrich's Periodicals Directory ©, U.S.A., Open J-Gage as well as in Cabell's Directories of Publishing Opportunities, U.S.A

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

Table-5.1
Year wise Aggregate Average RF for Poisson distribution estimates

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Average	15.168	40.71	66.56	83.5	85.576	70.414	49.622	30.604	18.116	9.508

Table 5.2
If we can arranged Ascending Order

Year	2016	2007	2015	2014	2008	2013	2009	2012	2010	2011
Average	9.508	15.168	18.116	30.604	40.71	49.622	66.56	70.414	83.5	85.576

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

$= 1.85$

1% l.o.f C.D $= 2.58 \times 1.85 = 4.77$

2016 2007 2015 2014 2008 2013 2009 2012 2010 2011

Above notation indicates that 2007-2015, 2009-2012, 2010-2011 years Average RF come under one category and 2016, 2014, 2008, 2013 year Average RF different category because there is no Significant Difference in average RF.

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

Table-5.3
Year wise Aggregate Average GWL for Poisson distribution estimates

Year	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016
Average	2.536	7.234	14.196	21.322	25.046	25.046	21.212	16.246	10.258	6.572

Table 5.4
If we can arranged Ascending Order

Year	2007	2016	2008	2015	2009	2014	2013	2010	2011	2012
Average	2.536	6.572	7.234	10.258	14.196	16.246	21.212	21.322	25.046	25.046

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

$= 1.15$

1% l.o.f C.D $= 2.58 \times 1.15 = 2.97$

2007 2016 2008 2015 2009 2014 2013 2010 2011 2012

Above notation indicates that 2016-2008, 2009-2014, 2013-2010, 2011-2012 years Average GWL come under one category and 2007, 2015 year Average GWL different category because there is no Significant Difference in average GWL.

ACKNOWLEDGEMENTS

The author expresses 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 author is profusely thankful to Late Retd. Prof. K.L.A.P. Sarma, in the Department of Statistics, S.K. University, Anantapuramu. Mr. C. Pothulaiah, Asst. Hydrogeologist, Ground Water and Water Audit Department and Mr. Muralimohan Reddy, Assistant Statistical Officer, Chief Planning Office, Anantapurmau. Mr. Madhav Reddy, Scientific Assistant PBO Anantapuramu Meteorological Department and also Mr. N.V. Subbaraju, Mr. Lanka Ramakrishna Prasad working as a Computer Faculty and finally Mr. Muralikrishna from Jaipur in Rajasthan (State) for his timely suggestions and useful discussions. Further I expression sincere thanks to Sri Krishnadevaraya University authorities for providing me the necessary facilities in the Department, to complete my Research work [5, 6 and 7].

AUTHOR PROFILE

Raju Sake is awarded PhD degree in Statistics under the Supervision of Retd. Prof. P. Mohammed Akhtar in the Department of Statistics, S.K. University, Anantapuramu and working as an Academic Consultant in the Department of Statistics, S.K. University, Anantapuramu.

REFERENCES

1. S. Raju, P. Mohammed Akhtar, *"Time Series Analysis on Rainfall and Ground Water Levels Data – A Case Study",* International Journal of Scientific Research in Mathematical and Statistical Sciences (IJSRMSS) Vol.6, Issue.1, pp.76-85, February(2019).
2. 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), pp. 01-06.
3. S. Raju, P. Mohammed Akhtar, *"Fitting Of Gompertz Model Between Rainfall And Ground Water Levels – A Case Study",* International Journal of Mathematics Trends and Technology (IJMTT) – Volume 65, Issue 7 – July (2019), pp. 85-93.

4. S. Raju, P. Mohammed Akhtar, *"Fitting of Logistic Model between Rainfall and Ground Water Levels – A Case Study",* Compliance Engineering Journal, Volume 10, Issue 9, pp. 114-122, September-2019.
5. S. Raju, P. Mohammed Akhtar, *"Fitting of Gompertz model between Rainfall and Ground water levels by using Partial sums method – A case study"*, Stochastic Modeling & Applications, Vol. 25 No.1 (January-June, 2021), MuK Publications and Distributions (INDIA). ISSN: 0972-3641.
6. S. Raju, *"Fitting of Modified Exponential Model between Rainfall and Ground Water Levels by Using Partial Sums Method – A Case Study",* Pensee Journal, Volume 51, Issue 03, pp. 647-655, March-2021.
7. S. Raju, "*Fitting of binomial distribution by using recurrence relation method between rainfall and ground water levels: A case study",* Journal of Mathematical Problems, Equations and Statistics (JMPES), Volume 1, Issue 02, pp. 03-08, 01-07-2020.
8. S. Raju, "*Fitting of Binomial Distribution between rainfall and ground water levels: A Case Study",* International Journal of Engineering, Science and Mathematics (IJESM), (ISSN: 2320-0294).