HS1.2.1 - Pathways & society transdisciplinary approaches towards solving the Unsolved Problems in Hydrology (UPH) : **EGU2020-4004**

Periodic occurrences of annual rainfalls in Eastern India

[UPH No. 9 (theme: *Variability of extremes*) and UPH No.19 (theme: *Modelling methods*)]

Dr. Subhabrata Panda

Department of Soil and Water Conservation,
Faculty of Agriculture, Bidhan Chandra Krishi Viswavidyalaya,
West Bengal, India. E-mail: subhabratapanda@gmail.com

HS1.2.1 - Pathways & society transdisciplinary approaches towards solving the Unsolved Problems in Hydrology (UPH) : **EGU2020-4004**

Salient Features of presentation

- **To examine the nature of changes in rainfalls from year to year for predicting groundwater recharge in unconfined aquifer regions in eastern India for sustainable use of groundwater in crop cultivation.**
- Usually those changes in annual rainfalls were periodic, with no regular increasing or decreasing trends, but with some types of cyclic changes having excesses and deficits about the mean.
- **With the use of long period annual rainfall data (since 1901 up to 1980) from nine raingauge stations in eastern India an attempt was made to predict annual rainfall through application of ARIMA model and Polynomial Regression.**
- Polynomial Regression was attempted because that has no bias towards periodicity.
- **Acceptability of either any of those two models was examined through *t*-test.**
- For most of the places both the models were found to be accepted. But in the predicted portion, polynomial regression was not working well because the predicted data from polynomial regression were beyond the range of observed annual rainfall data.
- **So, ARIMA model may be accepted for predicting annual rainfalls in eastern India with 12 years (included in that function) of cyclic period about the mean, that is nearly resembling a solar cycle of about 11 years.**
- Such conclusion may be drawn through minute analysis of more data and other better models for prediction for use of predicted annual rainfalls for planning of other purposes.

- Long period monthly rainfall data of nine raingauge stations throughout eastern India were collected from India Meteorological Department, Pune, India.
- Any missing monthly rainfall data were found out by taking average of monthly data of preceding and following years.
- Then Long Period nine annual rainfall data Series throughout eastern India were found out.

HS1.2.1 : **EGU2020-4004 :** Periodic occurrences of annual rainfalls in Eastern India

Table :Nine Raingauge Stations in eastern India with periods for collected data series

Nine raingauge stations throughout eastern India	Location		Data series for the years
	Lat.	Long.	
1. Aijawl (Mizoram)	23.7271	92.7176	1901 to 1965
2. Imphal (Manipur)	24.7829	93.8859	1901 to 1984
3. Guwahati (Assam)	26.1480	91.7314	1901 to 1986
4. Shillong (Meghalaya)	25.5669	91.8561	
5. Cherrapunji (Meghalaya)	25.2777	91.7265	
6. Cuttack (Odisha)	20.4625	85.8830	1901 to 1987
7. Patna (Bihar)	25.5818	85.0864	
8. Agartala (Tripura)	23.8903	91.2440	
9. Krishnanagar (West Bengal)	23.4058	88.4907	

Modelled Period and Predicted Period

Nine raingauge stations throughout eastern India	Data series for the years	Modelled Period	Predicted Period
1. Aijawl (Mizoram)	1901 to 1965	1901 to 1960	1961 to 1965
2. Imphal (Manipur)	1901 to 1984	1901 to 1980	1981 to 1984
3. Guwahati (Assam)	1901 to 1986		1981 to 1986
4. Shillong (Meghalaya)			
5. Cherrapunji (Meghalaya)			
6. Cuttack (Odisha)	1901 to 1987		1981 to 1987
7. Patna (Bihar)			
8. Agartala (Tripura)			
9. Krishnanagar (West Bengal)			

Modelled Period and Predicted Period

- modelled period – data for fitting a model

- **predicted period -** data for years left in the series after modelled period

for evaluation of the model for prediction of future rainfalls.

Modelled period: Analysis of annual rainfall series

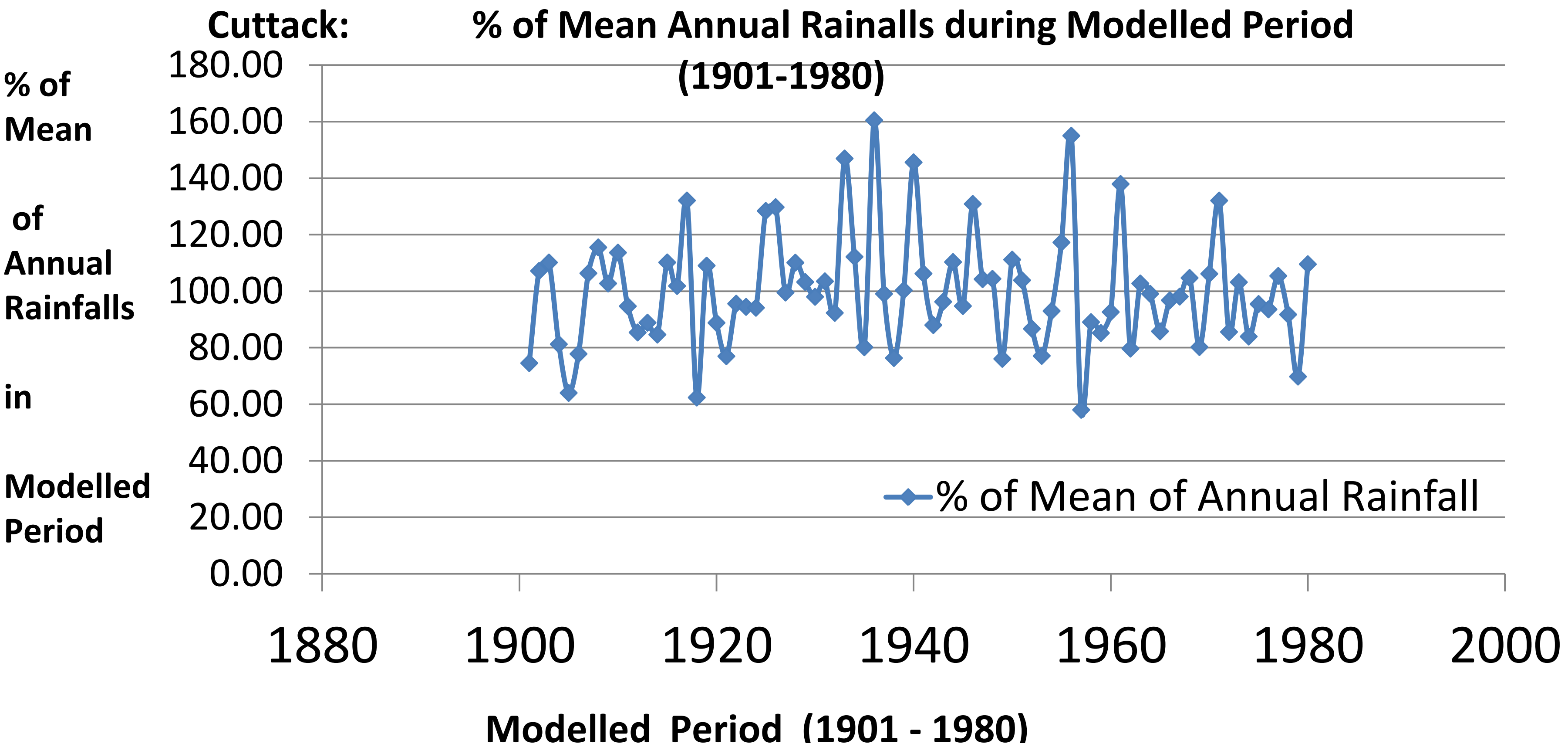
- Each annual rainfall series in the modelled period was first converted into percentage values of the mean annual rainfall

and then

- plotted against year, which showed the oscillations of the historigram about the mean line

(Tomlinson, 1987 for New Zealand rainfalls)

Cuttack: Oscillations of the historigram about the mean line



Modelled period: Analysis of annual rainfall series

- historigrams for all stations showed periodic nature of annual rainfalls throughout eastern India

- autoregressive integrated moving average (ARIMA) model was used to evolve a useful model for prediction of future rainfalls

1. Autoregressive integrated moving average (ARIMA) model

(*Clarke, 1973*)

1.1 The per cent annual rainfall series to be modelled for each station was analysed for periodicity following the ARIMA model (*Clarke, 1973*):

$$mt = \alpha + \beta\cos\left(\frac{2\pi t}{12}\right) + \gamma \mathrm{Sin}\left(\frac{2\pi t}{12}\right) + \epsilon_t \quad \ldots (1.1)$$

where mt = percentage of mean annual rainfall,

α = mean = $\sum \frac{mt}{n}$, expressed as 100 percent,

n = total number of years to be modelled, which is, here, divisible by 12,

t = particular number of a year 1, 2 ...n etc. from the beginning of the series,

$\beta = \left[\sum mt \left(\cos\frac{2\pi t}{12}\right)\right] / t /2$

$\gamma = \left[\sum mt \left(\cos\frac{2\pi t}{12}\right)\right] / t /2$

ϵ_t = Random Error

Random Error (ϵ_t) was minimised to obtain the accepted ARIMA equation (1.1) for use as Model to find out estimated values both for Modelled and Predicted Periods.

Autoregressive integrated moving average (ARIMA) model
(*Clarke, 1973)*

1.2 The 95 per cent confidence interval ($\widehat{\sigma\epsilon}^2$) or Delta for the estimated values from the equation (1.1) was found out as follows:

$$\widehat{\sigma\epsilon}^2 = \left\{ \sum_1^n (mt - \alpha)^2 - (\beta^2 + \gamma^2)\ n/2 \right\} (n - 3) \ldots\ldots (1.2)$$

1.3 ARIMA equation developed for Cuttack:

$$mt = 100.000 - 3.434\ \mathrm{Cos}\ \left(\frac{2\pi t}{12}\right) + 4.489\ \mathrm{Sin}\ \left(\frac{2\pi t}{12}\right) \ldots.\ (1.3)$$

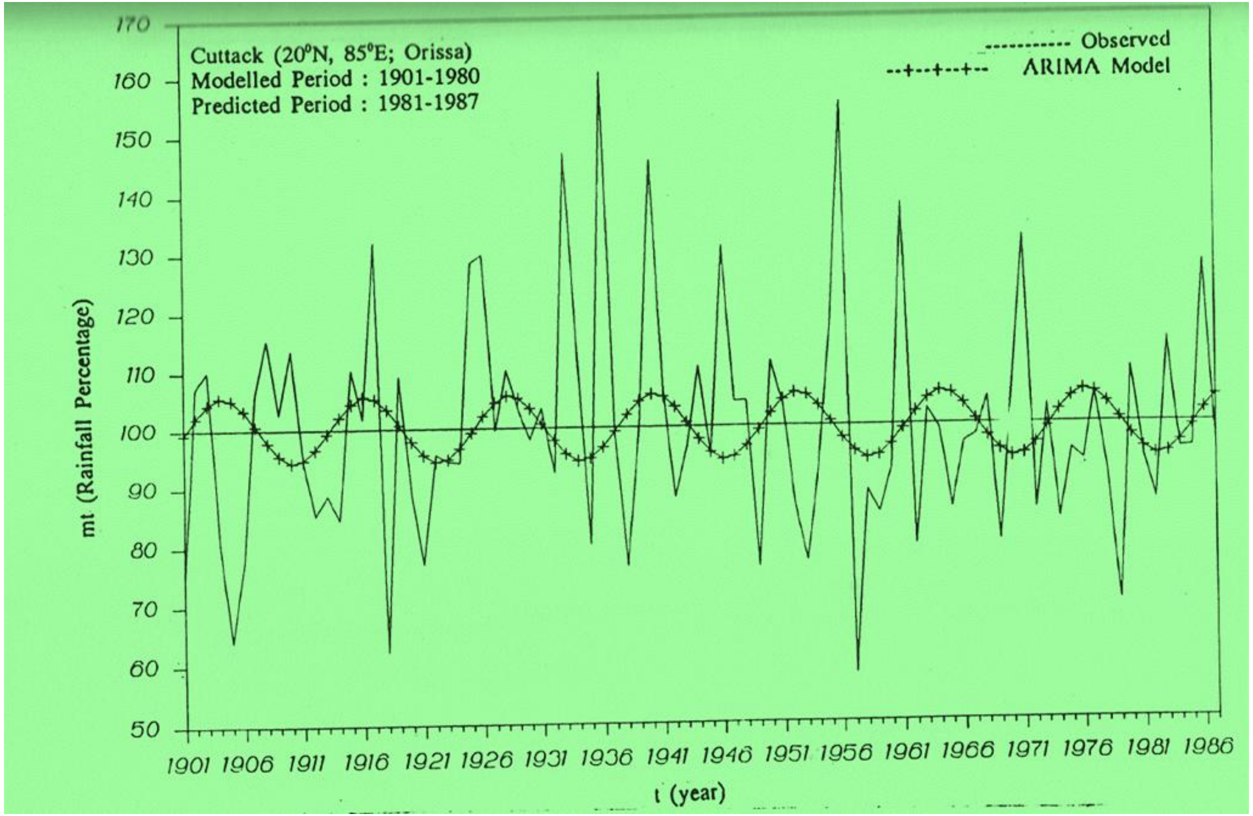
with 95 per cent confidence interval ($\widehat{\sigma\epsilon}^2$ or delta) of 20.419.

Autoregressive integrated moving average (ARIMA) model
(*Clarke, 1973)*

1.4 Variations in observed values of per cent annual rainfalls from those of estimated values from ARIMA model are visible from Fig 1 for Cuttack.

1.5 The variations of ARIMA model predicted mt from observed mt were calculated to judge the acceptance of the model (Table 1).

Variations in observed values from estimated values from ARIMA model



variations of ARIMA model predicted mt from observed mt

Table: Variation of ARIMA model predicted mt from observed mt at Cuttack

Year (t)	Observed mt (X mm)	Predicted mt (Y mm)	Variation of Y from X $\frac{X-Y}{X}$ *100per cent
1981	1435.40	1456.56	1.474
1982	1326.40	1439.55	8.531
1983	1745.30	1445.44	17.175
1984	1455.80	1472.66	1.158
1985	1459.60	1513.90	3.720
1986	1944.10	1558.13	19.853
1987	1455.50	1593.49	9.481

Modelled period: Analysis of annual rainfall series

ARIMA model Limitations

- ARIMA model was biased for periodicity due to inclusion of both the ‘sin’ and ‘cos’ functions and period length as 12.

Polynomial regression - application

- modelled data series were analysed for polynomial regression.

2.Polynomial regression

- 2.1 The periodicity in annual rainfall could also be studied through polynomial regression, because this regression actually helps us to find out the nature of the obtainable curve.

Polynomial regression

- 2.2 A polynomial regression of t is in the following form:

$mt = A + B_1t + B_2t^2 + ...+B_nt^n + \epsilon$.... (2.1)

Where A, B_1, B_2 ...Bn are coefficients of polynomial and are free from t; and n is a positive integer;

mt = per cent annual rainfall of a particular year (t);

t = concerned year, numbered as 1, 2 ....n;

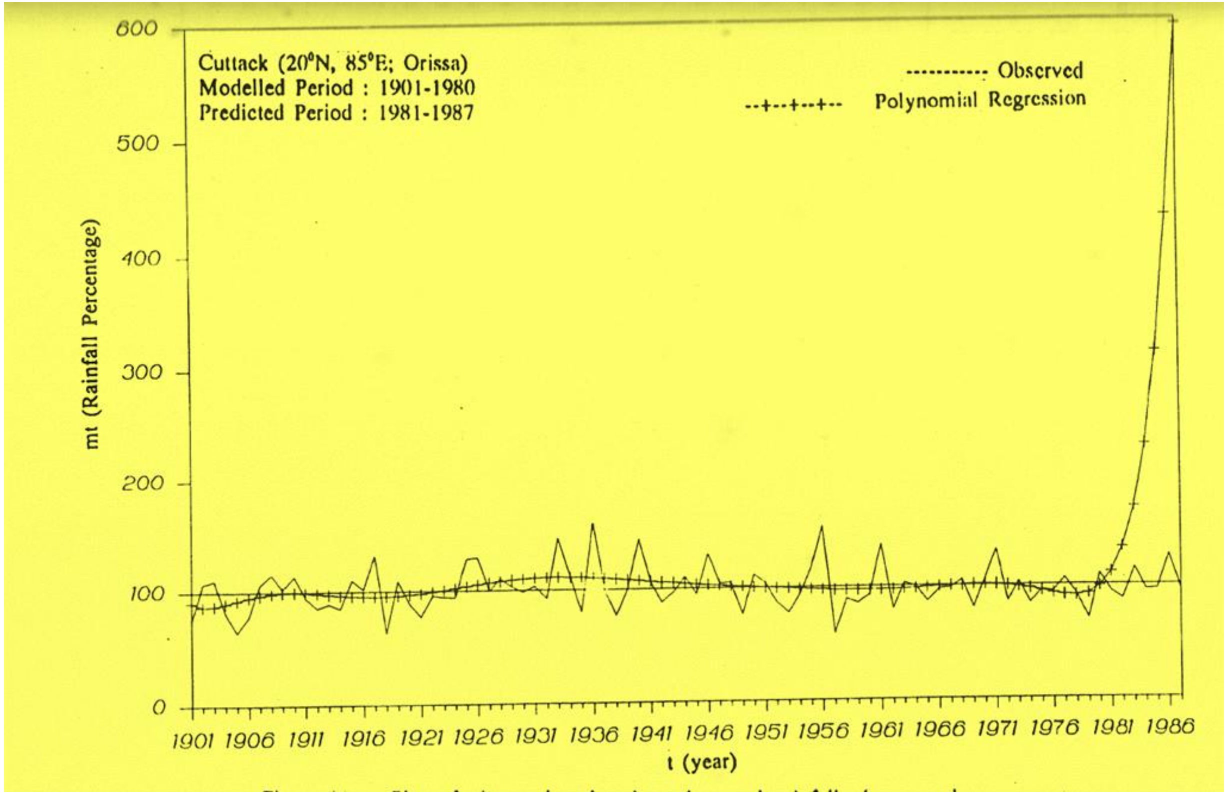
- ϵ, random error is minimised to obtain accepted degrees of polynomials as decided on the basis of analysis of variance (ANOVA).

2.3 Polynomial regression developed for Cuttack

- Polynomial Regression of Degree = 10
- Values of constants (10th order Polynomial)

```
 9.9557662055362OE+0001
-1.3255185357295OE+0001
 4.6451436034767 5E+0000
-6.4898138467924 6E-0001
 4.7398796191373 5E-0002
-2.0267494414802 9E-0003
 5.3598567878177OE-0005
-8.8978937730441 9E-0007
 9.0397420185743 7E-0009
-5.1404101870284 8E-0011
 1.2538122416772 3E-0013
```

Variations in observed values from estimated values from polynomial regression



Results:

Modelled period: Analysis of annual rainfall series

lengths of periods:

- **observed historigrams** –

  most cases - less than eight years

  and

  some cases - eight to 12 years

- **polynomial regressions**

  most cases -

  varied in between 8 to 12 years, 13 to 22 years and 23 to 37 years;

  and

  rare cases - 38 years and more

Modelled period: Analysis of annual rainfall series

3 Acceptance of either ARIMA model or polynomial regression decided on the basis of *t*-test.

3.1 t - test

$$t\text{-statistic} = \left|\frac{z_1 - z_1}{S.E.of\ (Z1 - Z2)}\right|$$

$$S.E.of\ (Z_1 - Z_2) = \sqrt{\frac{1}{n_1 - 3} + \frac{1}{n_2 - 3}}$$

Here, n_1 = n_2 = number of years modelled

for polynomial regression and ARIMA model respectively;

$Z_1 = \frac{1}{2} \log_e \frac{1+r_1}{1-r_1}$ and $Z_2 = \frac{1}{2} \log_e \frac{1+r_2}{1-r_2}$

Here, $r_1\, and\ r_2$ are correlation coefficients between years (t) and estimated per cent annual rainfalls (m_t) from polynomial regression and ARIMA model respectively.

3.2 t - test

If the value of $t-statistic$ is greater than 1.96 (i.e. the value of $t_{0.05}$ for infinite degrees of freedom, the difference between r_1 and r_2 is Then it could be concluded that out of these two models the second one i.e. ARIMA model is a better fit compared to the polynomial regression and this conclusion is accepted at 95 per cent probability level.

If the value of t-statistic, being less than that of 1.96, is proved to be non-significant, meaning thereby that the two estimates of correlation coefficient do not differ significantly.

3.3 *t* - test

Periods of Annual rainfall series	*t*	r(1) Polynomial Regression	r(2) ARIMA Model
Modelled Years	0.03757	-0.00249	-0.00854
Predicted Years	0.36788	0.95345	0.92291

For Cuttack values of for both the modelled and predicted periods, values of *t*-statistic being less than that of 1.96, is proved to be non-significant. So, two estimates of correlation coefficient do not differ significantly.

But for the predicted portion in Cuttack the ARIMA model predicted values vary within the range of observed mt.

Cuttack: Comparison between Polynomial regression and ARIMA model

- **MAXIMUM AND MINIMUM RAINFALLS**

Modelled Portion	Maximum	Year	Minimum	Year
Observed value (%)	160.48	1936	58.02	1957
Estimated(Polynomial)(%)	111.10	1933	87.13	1902
Estimated(ARIMA)(%)	105.61	1904	94.39	1910
Predicted Portion				
Observed value (mm)	1994.10	1986	1326.40	1982
Estimated(Polynomial)(mm)	9073.51	1987	1693.75	1981
Estimated(ARIMA)(mm)	1593.46	1987	1439.55	1982

HS1.2.1 : **EGU2020-4004 :** Periodic occurrences of annual rainfalls in Eastern India

Table :Nine Raingauge Stations in eastern India with *t* - values

Nine raingauge stations throughout eastern India	*t*- values		Remarks regarding Acceptance of ARIMA model in predicted period
	Modelled period	Predicted Period	
1. Aijawl (Mizoram)	2.58025	2.22338	ARIMA applicable
2. Imphal (Manipur)	2.65854	0.95250	predicted values vary within the range of observed mt
3. Guwahati (Assam)	0.52013	0.75714	
4. Shillong (Meghalaya)	0.16524	0.16837	
5. Cherrapunji (Meghalaya)	0.48867	0.23752	
6. Cuttack (Odisha)	0.03757	0.36788	
7. Patna (Bihar)	0.11364	1.37799	
8. Agartala (Tripura)	2.52464	4.16917	ARIMA applicable
9. Krishnanagar (West Bengal)	2.60582	0.84361	predicted values vary within the range of observed mt

Imphal: Comparison between Polynomial regression and ARIMA model

- **MAXIMUM AND MINIMUM RAINFALLS**

Modelled Portion	Maximum	Year	Minimum	Year
Observed value (%)	179.10	1966	62.49	1979
Estimated(Polynomial)(%)	123.70	1966	84.26	1976
Estimated(ARIMA)(%)	109.66	1905	90.34	1911
Predicted Portion				
Observed value (mm)	2555.70	1984	1077.10	1981
Estimated(Polynomial)(mm)	1487.89	1981	154.59	1984
Estimated(ARIMA)(mm)	1343.91	1981	1296.75	1983

Guwahati: Comparison between Polynomial regression and ARIMA model

- **MAXIMUM AND MINIMUM RAINFALLS**

Modelled Portion	Maximum	Year	Minimum	Year
Observed value (%)	151.70	1977	64.17	1944
Estimated(Polynomial)(%)	107.92	1974	92.94	1965
Estimated(ARIMA)(%)	110.15	1904	89.85	1910
Predicted Portion				
Observed value (mm)	1830.10	1983	1394.20	1981
Estimated(Polynomial)(mm)	12840.00	1986	1960.76	1981
Estimated(ARIMA)(mm)	1679.35	1986	1466.85	1982

Shillong: Comparison between Polynomial regression and ARIMA model

- **MAXIMUM AND MINIMUM RAINFALLS**

Modelled Portion	Maximum	Year	Minimum	Year
Observed value (%)	149.87	1902	64.13	1958
Estimated(Polynomial)(%)	150.14	1958	78.16	1978
Estimated(ARIMA)(%)	109.95	1903	90.05	1909
Predicted Portion				
Observed value (mm)	2370.10	1983	1754.40	1986
Estimated(Polynomial)(mm)	27555.44	1986	3147.00	1981
Estimated(ARIMA)(mm)	2391.88	1986	2006.32	1981

Cherrapunji: Comparison between Polynomial regression and ARIMA model

- **MAXIMUM AND MINIMUM RAINFALLS**

Modelled Portion	Maximum	Year	Minimum	Year
Observed value (%)	208.06	1974	51.45	1961
Estimated(Polynomial)(%)	125.59	1973	74.62	1964
Estimated(ARIMA)(%)	108.64	1904	91.36	1910
Predicted Portion				
Observed value (mm)	11811.30	1985	8696.90	1986
Estimated(Polynomial)(mm)	262450.65	1986	15732.66	1981
Estimated(ARIMA)(mm)	11422.18	1986	9995.79	1982

Cuttack: Comparison between Polynomial regression and ARIMA model

- **MAXIMUM AND MINIMUM RAINFALLS**

Modelled Portion	Maximum	Year	Minimum	Year
Observed value (%)	160.48	1936	58.02	1957
Estimated(Polynomial)(%)	111.10	1933	87.13	1902
Estimated(ARIMA)(%)	105.61	1904	94.39	1910
Predicted Portion				
Observed value (mm)	1994.10	1986	1326.40	1982
Estimated(Polynomial)(mm)	9073.51	1987	1693.75	1981
Estimated(ARIMA)(mm)	1593.46	1987	1439.55	1982

Patna: Comparison between Polynomial regression and ARIMA model

- **MAXIMUM AND MINIMUM RAINFALLS**

Modelled Portion	Maximum	Year	Minimum	Year
Observed value (%)	169.74	1918	52.23	1966
Estimated(Polynomial)(%)	127.83	1977	59.48	1901
Estimated(ARIMA)(%)	114.76	1904	85.24	1910
Predicted Portion				
Observed value (mm)	1885.60	1987	699.00	1982
Estimated(Polynomial)(mm)	2577.80	1987	838.95	1984
Estimated(ARIMA)(mm)	1299.23	1987	983.87	1982

Krishnanagar: Comparison between Polynomial regression and ARIMA model

- **MAXIMUM AND MINIMUM RAINFALLS**

Modelled Portion	Maximum	Year	Minimum	Year
Observed value (%)	161.21	1905	37.41	1979
Estimated(Polynomial)(%)	120.96	1910	47.03	1980
Estimated(ARIMA)(%)	108.26	1904	91.74	1910
Predicted Portion				
Observed value (mm)	1560.00	1984	606.10	1982
Estimated(Polynomial)(mm)	4570.37	1987	421.01	1982
Estimated(ARIMA)(mm)	1528.77	1987	1304.96	1982

Conclusions:

- Considering all limitations in the observed data and 95% confidence interval for ARIMA model,
- a particular amount of annual rainfall occurred at about 12 years
- (i.e. almost resembling a Solar Cycle of about 11 years) and that needs minute analysis of more observed data.
- Recurrence of flood and drought years can be predicted from such analysis

and also by following

- probability analysis of excess and deficit runs of annual rainfalls

(Panda *et al*., 1996).

HS1.2.1 : EGU2020-4004 : Periodic occurrences of annual rainfalls in Eastern India

References:

Clarke, R.T.1973. Mathematical models in hydrology. FAO Irrigation and Drainage Paper No. 19. FAO of the United Nations, Rome. pp.101-108.

Panda, S.; Datta, D.K. and Das, M.N. (1996). Prediction of drought and flood years in Eastern India using length of runs of annual rainfall. J. Soil Wat. Conserv. India. 40(3&4):184-191. (https://www.academia.edu/15034719/Prediction_of_drought_and_flood_years_in_eastern_India using_length_of_runs_of_annual_rainfall)

Tomlinson, A.I. (1987). Wet and dry years – seven years on. Soil & Water. Winter 1987: 8-9. ISSN 0038-0695

Thank you