Precipitation Prediction by Pattern Recognition and Statistical Analysis of 110 Years

Shailesh Gupta
B.Tech Maharaja Agrasen Institute of Technology,
GGSIPU, New Delhi, India
shaileshgupta94@gmail.com

Shaurya Vardhan
B.Tech Maharaja Agrasen Institute of Technology,
GGSIPU, New Delhi, India
shauryavardhan1@gmail.com

Abstract: ***This paper presents the Pattern in rainfall observed over the last 110 years in the month of March, in NCR Region, India. The present forecasting system has timely failed to predict the certainty of rainfall in the month of March, which has led to the huge destruction of crops, which further sometimes causes farmers to attempt suicide as their only source of income is destroyed. Thus, shortage of grains leads to hiking in prices of the food. Therefore, if the timely warning is issued about the uncertain rainfall in March, the loss of crops and in turn lives of farmers can be saved. Our proposed Prediction system predicts rainfall well in advance as it does not consider any other factors on which precipitation depends. It predicts on the basis of pattern and trends in rainfall over the last century, which has been proved correct in each decade of the 20th century.***

Keywords: Pattern Recognition, Statistical Analysis, Precipitation, Western Cyclonic Disturbances, Late Sowing.

1. Introduction

Western disturbances originate in the Mediterranean region. A high-pressure-area over Ukraine and neighbourhood consolidates, causing the intrusion of cold air from polar regions towards an area of relatively warmer air with high moisture. This generates favourable conditions for cyclogenesis in the upper atmosphere, which promotes the formation of an eastward-moving extratropical depression. Traveling at speeds up to 12 m/s (43 km/h; 27 mph), the disturbance moves towards the Indian subcontinent until the Himalayas inhibits its development, upon which the depression rapidly weakens.

The concept of Western Cyclonic Disturbances justifies the late winter rainfall in the month of Jan/Feb. But, sometimes due to late Western Cyclonic Disturbances, it rains in the month of March, the March rainfall is never good for the crops as they are very near their Harvesting season. If the farmers are warned well before time about uncertain March rainfall then they can use the concept of late sowing, where they sow the seeds in the month of December rather than Oct-Nov. In late sowing, the harvesting takes place in the month of May, so any rainfall in March will actually benefit the crops rather than Destroying them completely.

2. Pattern Found in the Data

The data for the amount of monthly rainfall from 1901-2010 was obtained from www.opengovdata.in.

On Rigorous Analysis of the data, the pattern we found out was that for every 10 years, the probability of occurrence of rainfall was greater than 15mm in the month of March for the median years of the corresponding decade. For example: in the years of the decade 2001-2010, the probability of it raining more than 15 mm in the month of March is most in the years 2004,2005,2006,2007 and 2008.

So, spreading awareness in these 5 out of 10 years of the decade can prevent large scale devastation of crops and eventually human life.

3. Existing and Proposed System

3.1 Existing System

The current weather forecasting system fails to predict and tell well in advance about the rainfall for the month of March. Though the Metrological Department of India gives quite good measures of the rainfall, it lacks in forecasting rainfall in the Month of March, thus causing widespread destruction of crops.

Furthermore, the Metrological Department of India does not acknowledge the farmers about the rainfall, thus leading to the mass destruction.

Thus, our proposed Prediction system predicts rainfall well in advance as it does not consider any other factors on which precipitation depends. It predicts on the basis of pattern and trends in rainfall over the last century, which has been proved correct in each decade of the 20th century.

Limitations of the existing system

- Failed to Predict the certainty of the amount of rainfall in the month of march. Thus, causing widespread destruction of crops.
- Does not acknowledge farmers about the rainfall.
- Too complex to be understood by the illiterate and common farmers of the country.

3.2 Proposed System

The Proposed System is based on Numerical Weather Prediction.

Numerical weather prediction (NWP) uses mathematical models of the atmosphere and oceans to predict the weather based on current weather conditions.

Mathematical models based on the same physical principles can be used to generate either short-term weather forecasts or longer-term climate predictions; the latter is widely applied for understanding and projecting climate change.

The horizontal domain of a model is either global, covering the entire Earth, or regional (we have used regional horizontal domain), covering only part of the Earth. Regional models (also known as limited-area models, or LAMs) allow for the use of finer grid spacing than global models because the available computational resources are focused on a specific area instead of being spread over the globe. Following are the salient features of our model:

- The simplicity of WeatherX is that it doesn't analyze too many parameters.
- It only Focuses on Trend Analysis of the past century.
- The system is tested at 90% Accuracy in the median years and 60-70% accuracy in general.

The new model, for now, takes into the consideration the amount of precipitation in the previous decades. After analyzing 110 years of precipitation data, taken from the Metrological Department of India, we come to find a unique pattern. This pattern shows that the probability of rainfall happening in the month of March is highest for the Median years of the corresponding decade.

3.3 Data Analysis

In the next section of the research paper, you will find out various images and tables which will help you visualize the pattern found in the data. As you can see, for the graph of each decade there is a spike around the median years which shows the anomaly that there are more chances of raining than the average in corresponding years of the decade.

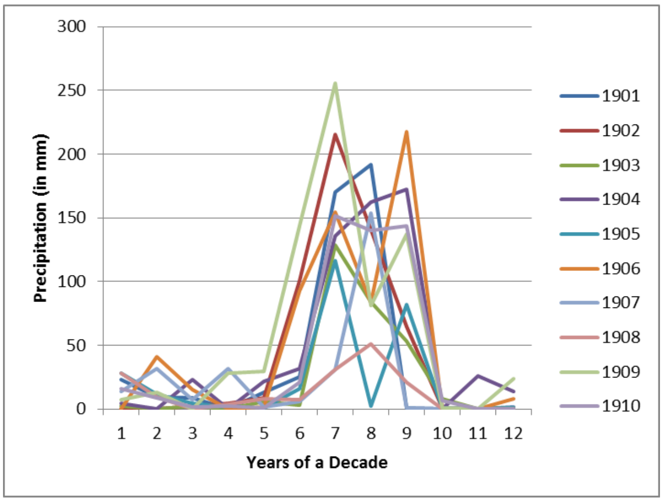
3.4 Tables

	Truth (mm)	Observed (mm)	Error
1.	10.8375	10.7	0.01262
2.	11.5154	11.1	0.03607
3.	4.28163	8.7	-1.03193
4.	**10.2325**	**9.28**	**0.09309**
5.	**11.6659**	**13.1**	**-0.12293**
6.	**14.4443**	**17.4**	**-0.20462**
7.	**16.6475**	**18.8**	**-0.1293**
8.	**15.7228**	**16.2**	**-0.03035**
9.	5.19081	12.7	-1.44663
10.	8.33054	18.8	-1.25675

Truth and Observed rainfall in mm. The error is a standard deviation from the original rainfall pattern.

It can be clearly seen from the table that error in the median years (Highlighted), is between -1 and +1, it is a significant discovery. For Visual representation, refer the ERROR graph below.

1)1901-1910



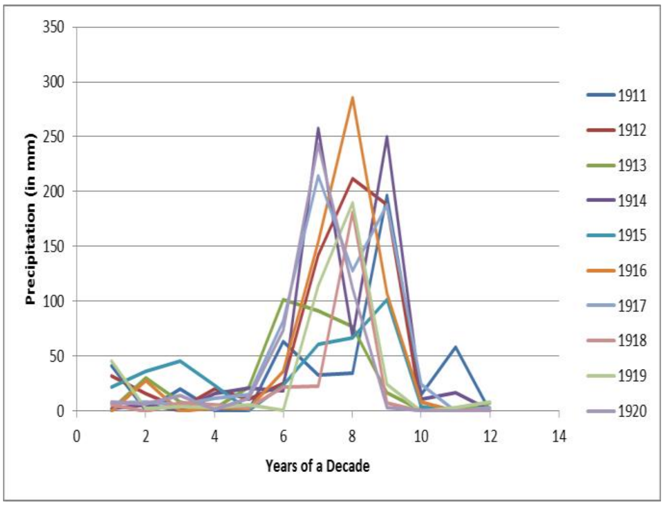
3.5 Figures

<u>Trend Analysis for all decades</u>:

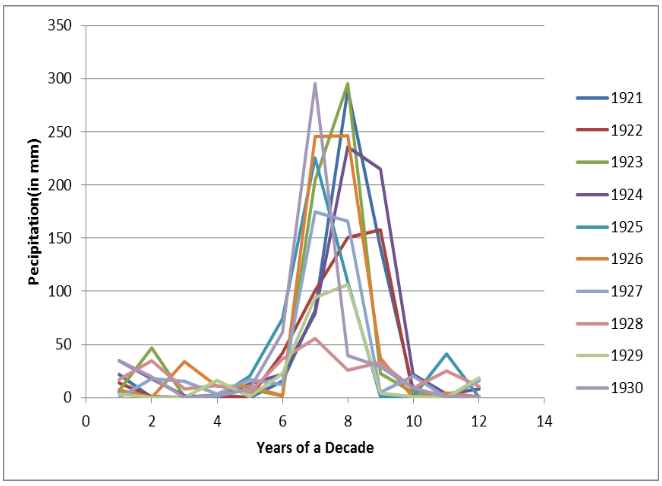
In the following graphs, the x axis represents years of a decade from 1 to 10, and y axis represents the average precipitation in the month of March
In that particular year.

Note: Precipitation data for the years 1951-60 has been deleted in order to support and improve our equation since it was the only anomaly in the century.

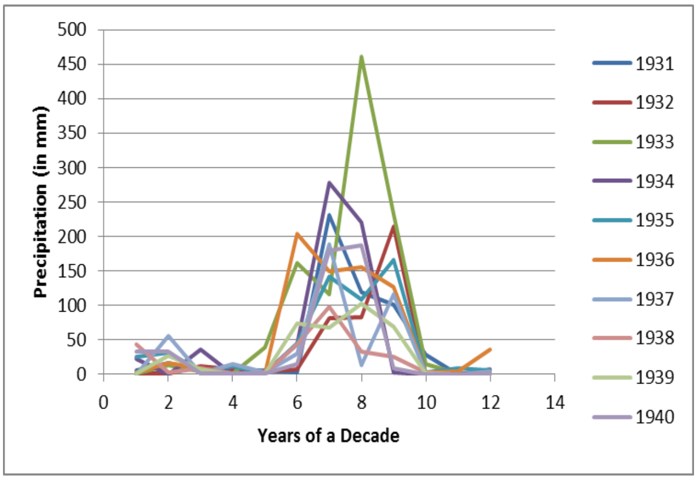
2)1911-1920



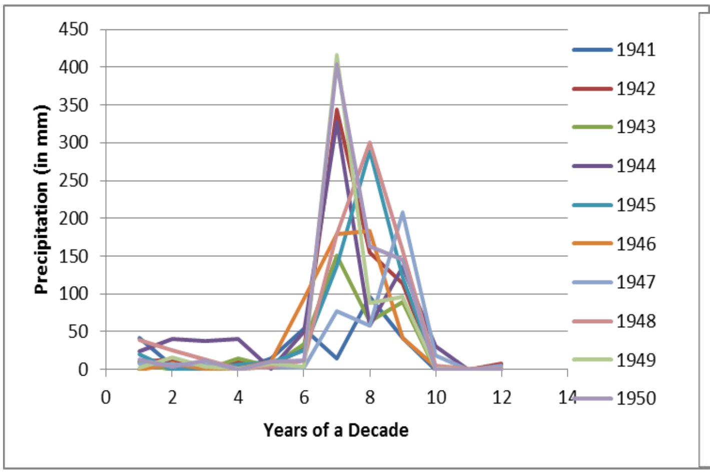
3)1921-1930



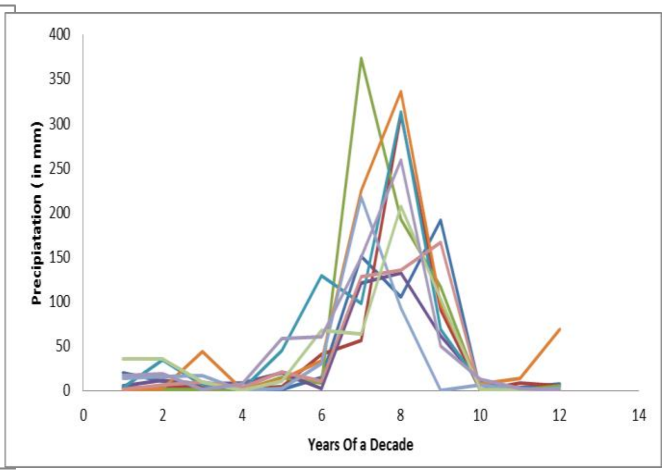
4)1931-1940



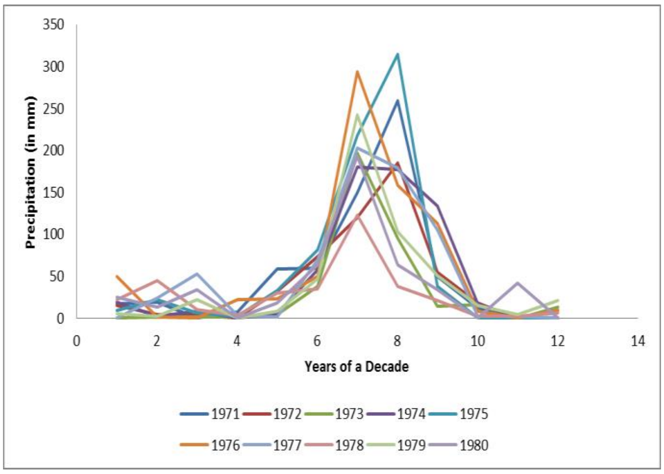
5)1941-1950



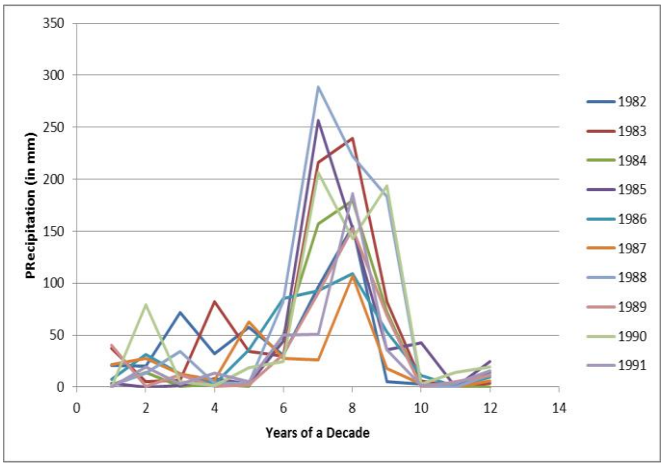
6)1961-1970



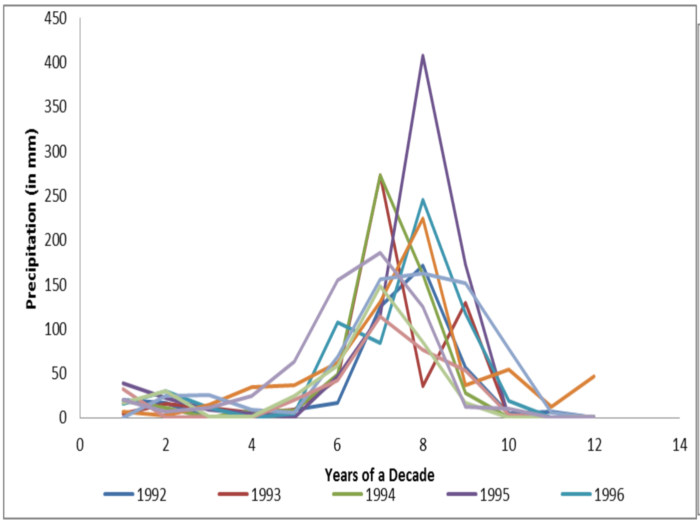
7) 1971-1980



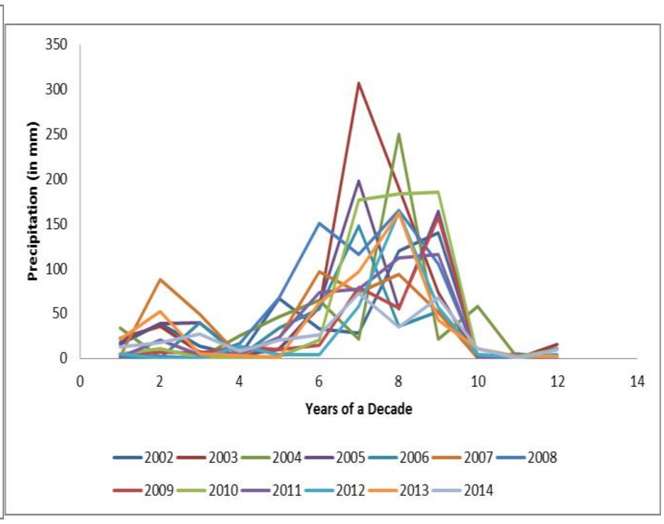
8) 1982-1991



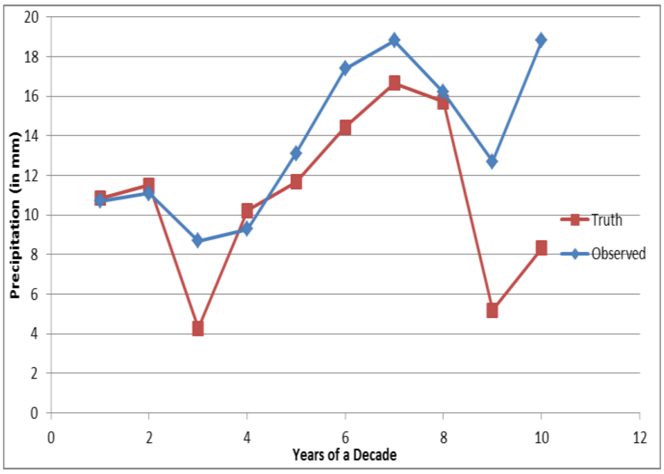
9)1992-2001



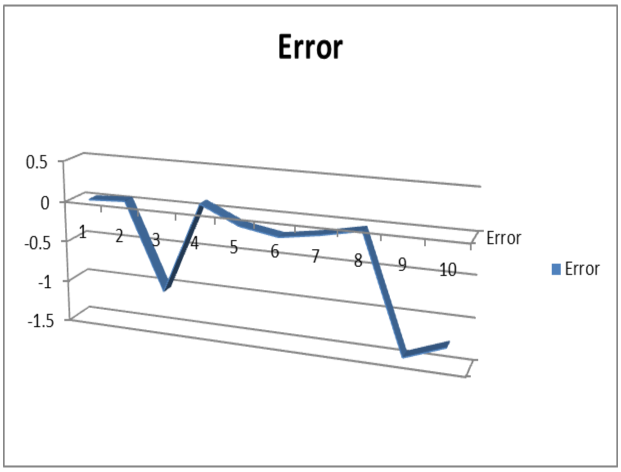
10) 2002-2014



Truth V/S Observed Values



Error: The following graph shows Error vs Years of a Decade.



4. Equation

(1) The equation that represents the pattern found in the data is a fifth order polynomial equation.
(2) Here x is the year of the decade, for example, if you want to enter 2014, just input 4 as the value of x.
(3) Y is the amount of precipitation in the month of March of that year. Y is your output.

$$Y=a*x^5 + b*x^4 + c*x^3 + d*x^2 + e*x + f$$

- *a=0.01991*
- *b=-0.53387*
- *c=5.1258*
- *d=-21.1397*
- *e=36.3208*
- *f= -11.0364*

5. Future enhancements

.

- We can search information more promptly and we can use the results conveniently.
- Ease of access.
- Faster results.
- Interactive design of the interface.
- Analyze more parameters such as Temperature, wind speed and more.

6. Conclusion

- At present WeatherX only predicts the amount of rainfall in the month of March for Delhi/NCR region. But later on, it can be developed into a fully functional system that can predict rainfall for any place across the country and for any crop.
- It can also help with Flood Forecasting in future.
- Flash flood analysis and prediction models are now coming into widespread use in some countries. Flash flood models represent the topography, surface conditions, soil moisture and stream networks of flood plains and basins, and they rely on observations or model predictions of precipitation rates and duration to make short-term forecasts of flooding.
- It can help in suggesting different types of cropping patterns for different types of crops such as rabbi crops and Kharif crops.

7. References

[1] Localized Precipitation Forecasts from a Numerical Weather Prediction Model Using Artificial Neural Networks
[2] The Intra-seasonal Oscillation and the Energetics of Summertime Tropical Western North Pacific Synoptic-Scale Disturbances
[3] A Statistical-Topographic Model for Mapping Climatological Precipitation over Mountainous Terrain
[4] https://en.wikipedia.org/wiki/Western_Disturbance
[5] www.opengovdata.in

8. Author Profile

Both the authors are recent graduates (Class of 2017) of Maharaja Agrasen Institute of technology, New Delhi, IT branch.