A STATISTICAL MODEL FOR PREDICTION OF ARIAL RAINFALL OVER BARAKAR CATCHMENT

1. In this work, the factor we consider the most for forecasting quantitative rainfall over an area is the synoptic meteorological situation prevalent near and over the area and associated rainfall with it in past events.

In India, weather forecast which is valid for 24 hours and 48 hours are issued on the basis of weather situations of 0300 UTC and 1200 UTC observations. Many researchers have attempted to make a statistical study of the rainfall and associate it with synoptic meteorological situation.

In this model we are mainly following the work of Basu (1989) and trying to enhance the model for Barakar Catchment following Wilks (1995).

The statistical theory of association in forecasting the rainfall amount by utilizing the prevailing synoptic meteorological condition has been studied for Barakar catchment area in monsoon season period from 1997 to 2002. We are trying to correlate the meteorological condition i.e., how the atmospheric pressure and trough line has influenced the rainfall in catchment area. We have categorized different synoptic situation under column \((B_j, j = 1 \text{ to } 10)\) etc. and amount of rainfall in row \((A_i, i = 1 \text{ to } 3)\). The above columns and rows give rise to a generic contingency \(3 \times 10\) table which is the basic platform of our analysis. With this table we form a number \(2 \times 2\) contingency table and calculate the respective Yule and Tschuprow’s coefficient.

2. Method of the model (i) Contingency table and association of attributes: The relationship of any kind between two attributes \(A_i\) and \(B_j\) and their association as suggested by Yule (Wilks, 1995) has been applied here. The statistical frequency distribution \(o_{ij}\) for the rainfall ranges responsible for the particular synoptic condition has been arranged in two way contingency table. The frequencies of such distribution of three different ranges of rainfall amount as rows \((A_i, i = 1 \text{ to } 3)\) occurring for ten categories of synoptic weather situation as columns \((B_j, j = 1 \text{ to } 10)\) has been set into \(3 \times 10\) contingency table with \((10-1) \times (3-1) = 18\) degrees of freedom. The expected frequencies \(e_{ij}\) in each cell has been found out by:

\[
e_{ij} = \frac{(A_i) \times (B_j)}{N}
\]  

where \(A_i\) and \(B_j\) are the total frequencies of \(i\text{th}\) row and \(j\text{th}\) column respectively. \(N\) is the total number of frequencies in the contingency table. The null hypothesis of independence between any of the two attributes \(A_i\) and \(B_j\) is tested by \(\chi^2\) statistic given by:

\[
\chi^2 = \sum \sum (o_{ij} - e_{ij})^2
\]  

Now, to measure the association of each of the particular range of rainfall for the particular synoptic condition, the contingency table has been reduced to a number of \(2 \times 2\) tables with one degree freedom. Yate’s correction \(\pm 0.5\) is to be added with the frequencies for continuity in the \(2 \times 2\) tables in such a way that the sum of any row or column remain unchanged. This correction was necessary in those cell entries where \(o_{ij} < 5\).

(ii) Measure of association: The \(2 \times 2\) tables after Yate’s correction ( wherever necessary ) reduce to a form:

\[
\begin{array}{ccc}
 & B & \text{not } B & \text{Total} \\
A & a & b & a+b \\
\text{not } A & c & d & c+d \\
\text{Total} & a+c & b+d & N \\
\end{array}
\]

The association of the variate of the above \(2 \times 2\) tables by Yule’s coefficient \((Y)\) is given by

\[
Y = \frac{(ad-bc)}{(ad+bc)}
\]

Where,

(a) \(Y = 0\) implies independence of \(A\) and \(B\).
(b) \(Y = 1\) implies complete association of \(A\) and \(B\).
(c) \(Y = -1\) implies complete dissociation.

### TABLE 1

<table>
<thead>
<tr>
<th>Rainfall range in mm</th>
<th>Synoptic Meteorological Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(B_1)</td>
</tr>
<tr>
<td>(A_1)</td>
<td>10</td>
</tr>
<tr>
<td>(A_2)</td>
<td>16</td>
</tr>
<tr>
<td>(A_3)</td>
<td>10</td>
</tr>
<tr>
<td>Total</td>
<td>36</td>
</tr>
</tbody>
</table>

The frequency of average aerial rainfall in different categories of synoptic weather conditions over Barakar basin from 1997 to 2002.
The values of Yule’s coefficient (Y) of each 2 × 2 tables were put into Table 2.

Using the following formula $T = (\chi^2 / N)^{\frac{1}{2}}$ (4)

Tschuprow’s coefficient (T) of the above 2 × 2 table for degree of associations or dissociations were calculated and put into Table 3.

In the Eqn. (4) $\chi^2$ - statistic represents the sum of $\chi^2$ - statistics calculated from the above 2 × 2 table and N is the total number of frequencies.

3. Significant synoptic feature categorization -

The study of the corresponding synoptic weather situations responsible for the significant amount of rainfall over Barakar catchment area based on the inferences from DVC Met. Unit, RMC, Kolkata and weather charts are categorized into ten classes and are listed below:

- $B_1$ -> Low pressure area (LOPAR) covering wholly or partly the basin area.
- $B_2$ -> LOPAR lying in the neighbourhood, i.e., outside the basin area which may occur over Bihar plains, Bihar plateau, Bangladesh, north Orissa or east Madhya Pradesh.
- $B_3$ -> LOPAR / DEPRESSION in north west Bay of Bengal and neighbourhood and areas covering partly or near the basin area.
- $B_4$ -> Well marked LOPAR / DEPRESSION covering partly /wholly or near the basin area with associated CYCIR (Cyclonic Circulation) extending upto mid tropospheric level.
- $B_5$ -> LOPAR near the basin with axis of trough line passing near south of basin area.
- $B_6$ -> The axis of the trough line passing through the basin area.
- $B_7$ -> The axis of the trough line passing through north of basin area.
- $B_8$ -> CYCIR within basin area and trough within basin area.
- $B_9$ -> CYCIR within basin area and trough outside basin area.
- $B_{10}$ -> CYCIR outside basin area and trough outside basin area.

### TABLE 2

<table>
<thead>
<tr>
<th>Rainfall range in mm</th>
<th>Synoptic Meteorological Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$B_1$</td>
</tr>
<tr>
<td>$A_1$</td>
<td>-0.698</td>
</tr>
<tr>
<td>$A_2$</td>
<td>0.329</td>
</tr>
<tr>
<td>$A_3$</td>
<td>0.856</td>
</tr>
</tbody>
</table>

### TABLE 3

<table>
<thead>
<tr>
<th>Rainfall range in mm</th>
<th>Synoptic Meteorological Conditions</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$B_1$</td>
</tr>
<tr>
<td>$A_1$</td>
<td>0.231</td>
</tr>
<tr>
<td>$A_2$</td>
<td>0.093</td>
</tr>
<tr>
<td>$A_3$</td>
<td>0.313</td>
</tr>
</tbody>
</table>
4. Data and method of compilation - The daily rainfall data based on 0300 UTC observations of some selected stations over Barakar basin area were plotted in the basin map. The area lies between longitude 85° 05' E and 86° 55' E and latitude 23° 40' N and 24° 30' N. The daily average rainfall (AR) over the basin area has been calculated by isohyetal analysis method. Those daily AR were considered here for study and were categorized into three different classes, namely

- $A_1$ for 11-25 mm rainfall
- $A_2$ for 26-50 mm rainfall
- $A_3$ for 51 and above rainfall

The frequencies and probability of occurrences of such rainfall ranges influenced by the particular class of synoptic weather situation are arranged in $3 \times 10$ contingency (Table 1). The expected frequencies were found out by equation (1) from the observed frequencies of occurrence.

The $\chi^2$ - test with the appropriate degrees of freedom for dependency of the system was tested. To estimate the association with different ranges of AR for different synoptic situations the original $3 \times 10$ contingency table gets reduced into a number of $2 \times 2$ tables. The Yule’s coefficient of association and Tschuprow’s coefficient are found out by Equations (3) and (4) for each of the $2 \times 2$ tables, after applying Yate’s correction wherever necessary.

5. Discussion - The measure of intensities of association by Yule’s coefficient ($Y$) and by Tschuprow’s coefficient ($T$) for each of the rainfall ranges for particular synoptic weather situations, considering $2 \times 2$ tables calculated by Equations (3) and (4) respectively and is given in Table 2 and Table 3 respectively and the following facts noted.

(i) The AR ranges 51 mm and above has the highest association of about 85% for the category situation $B_1$, and corresponding Tschuprow’s coefficient is 0.313.

(ii) The AR ranges 26-50 mm and above has the highest association of about 62% for the category situation $B_3$, and corresponding Tschuprow’s coefficient is 0.130.

(iii) The AR ranges 11-25 mm and above has the highest association of about 69% for the category situation $B_3$, and corresponding Tschuprow’s coefficient is 0.156.

From the Tables 2 and 3, seeing the Yule’s and Tschuprow’s coefficient one can get a notion of the possibility of occurrence of AR ranges over Barakar Catchment area with varied synoptic situation.

6. Application in forecasting - The categorized data of daily AAR in different synoptic weather situations as displayed in Table 1 and the study of their association or dissociation over any area (here it is Barakar basin) displayed in Tables 2 and 3, is very much useful in predicting rainfall over that area. So, this statistical synoptic model derived from association method could be important guide in forecasting the daily AR over the basin area.

7. Conclusion - The present study based on the statistical association method applied for prediction of average rainfall with reference to ten categories of synoptic situations (as labeled $B_j$, $j = 1$ to 10) over Barakar basin area reveals the following:

(i) The association of rainfall range 51 mm and above is maximum with the synoptic situation of category $B_1$, i.e., “Low pressure area (LOPAR) covering wholly or partly the basin area”. The leastassociation of this range is with the synoptic situation of category $B_2$, i.e., “LOPAR lying in the neighbourhood i.e., outside the basin area which may occur over Bihar plains, Bihar plateau, Bangladesh, north Orissa or east Madhya Pradesh”.

(ii) The association of rainfall range 26-50 mm is maximum with the synoptic situation of category $B_5$, i.e., “LOPAR near the basin with axis of trough line passing near south of basin area”.

(iii) The association of rainfall range 11-25 mm is maximum with synoptic situations of category $B_3$, i.e., “LOPAR near the basin with axis of trough line passing near south of basin area”.

The above analysis is expected to be useful in providing forecasts of average rainfall under one of the synoptic situations categorized here. Further detailed studies are required for taking position, movement and intensity of the weather systems along with movement of the earth-moon system, which will give rise to better insight for catchment area rainfall forecasts. In future study with more parameters as stated above, we may employ rule based expert system following Konar, (2000) or Karman filtering technique following Hunt et al. (2006).
KILLER TORNADOES DURING HIGH TIDE PERIOD

1. Solar - lunar gravitational influence produces Ocean tides (spring and neap) during different phases of the Moon. Values of both high tides of the semi-diurnal tides (two high waters and two low waters each day) start decreasing normally from 4th or 5th day, continue decreasing till 8th or 9th day, again start increasing from 10th or 11th or 12th or 13th day and continue increasing till 1st or 2nd or 3rd day of waning and waxing period of the lunar month (15 days each of Krishna paksha and Shukla paksha of Hindu calendar). Ocean tides exhibit the intensity of the gravitational force of the Moon on a particular day. It has been observed that more killer tornadoes (for USA: fatalities 39 or more or massive destruction reported or intensity: EF 3 or more on enhanced Fujita scale; for Asian countries: fatalities 28 or more or massive destruction) develop during high tide period (10th day to 3rd day in both fortnights) than low tide period (4th to 9th day in both fortnights) under favorable synoptic situations.

2. The stretch of the United States, which runs from Kansas through Oklahoma to Missouri, is known as “Tornado Alley”: a section in which more tornadoes strike than in any other place in the world. No area is more favorable to their formations than the continental plane of North America (Lynch 2002). Ramasastry (1984) has mentioned that the most preferred regions of their occurrence in India are Assam and adjoining states, West Bengal, Gangetic planes, Punjab and Haryana. He has further mentioned that 72% of the reported tornadoes have occurred in northeast India. When super cell thunderstorm struck Bihar, West Bengal and Bangladesh on 13th April 2010, it was 14th lunar day of waning period (Krishna paksha Chaturdashi of Hindu calendar). Total fatalities due to these were 120 and severe destructions have been reported (News papers reported tornado) along its path. Longest distance covered by the super cell thunderstorm on 13th April did not suggest a tornado.

3. Year wise tornadoes data for 66 years period (1944-2010) have been collected from Google web site. Shri Venketeshwar Shatabadi Panchangam (hundred years Almanac: 1944 to 2044 AD) has been used to identify lunar day for reported tornadoes. Tide tables for 2009 and 2010 (prepared by Survey of India) have been used.

4. Ramasastry (1984) has mentioned that thunderstorms usually develop in clusters consisting of several thunderstorm cells, each about 4 km across. They tend to move 20° to the right of mid tropospheric direction. During this motion new cells develop to the