Chapter 1

INTRODUCTION
1.1. Introduction to Cloud Seeding Experiments:

Water is the most important and crucial factor for the development of any area/nation. India is a developing country, mainly depending on agricultural sector, where water plays an important role. Thus, the progress of this nation depends mainly on agriculture. To improve agricultural returns, sufficient water resources are to be developed. Hence, there is a necessity to concentrate on appropriate management of various 'Water resources' available. Basically, water resources are of two types namely:

i) Surface water resources and

ii) Ground water resources. [18]

In agriculture sector, both resources play a crucial role. Now-a-days, agricultural bore-wells are playing vital role to harness the groundwater resources in areas where surface water resources are scanty [11]. Further, to minimize the water wastage and to utilize the limited resources judiciously, new agricultural techniques like i) Drip Irrigation, ii) Water Sprinkling and so on have come into practice. Water resources management has become vital to get optimum returns from the agricultural sector. Ground water and Surface water gets replenished through rainfall and rivers [11].

Anantapur, a district of Andhra Pradesh is taken up as our study area for this Thesis work. Unfortunately, Anantapur district is not only one of the drought prone areas but also has no perennial rivers flowing into it. Due to the lack of perennial water sources, the farmers of this district are forced to depend mainly on limited rain water for providing irrigation to their agricultural crops. Here, the rain fall comes usually in two spells of monsoon namely:
1) **South - West monsoon (June, July and August months)** and

2) **North - East monsoon (September, October and November months)**

[40(i), (iii), (iv) & (v)].

Further, major portions of this district have been in a rain-shadow zone for several years continuously [40(ii), (vi), & (vii)].

In this district, nearly 60% of people are farmers and agricultural labourers, whose livelihood will be comfortable only if there are good rains in these areas, otherwise, their lives will be miserable. For the last few years, rainfall has been very low in this district and many farmers are in long pending debts, ultimately leading to migration to other areas for their mere survival. The Government of Andhra Pradesh, many scientists, researchers, intellectuals, industrialists, statesman related to agricultural industry are gripped with this problem and looking for different ways and means, to improve water resources of this district using latest scientific techniques [40].

On **19th October, 2006** a news item in 'The Hindu' newspaper indicates, **acute water crisis in the district of Anantapur.** Many areas required transportation of drinking water through tankers. The impact of fast depleting groundwater table has a telling effect on the farming community as a few farmers have ended their lives due to drying up of farm borewells in the previous few months. Poor rainfall has enabled cultivation of only 60 per cent of the land normally cultivated in the khariff season, according to the official estimates. Even the Rural Water Supply Department indicated that 530 of the
14,383 hand pumps under their purview had gone dry. Water Table is going down day-by-day, in 18 mandals it has been recorded between 15-20 meters below surface, between 10-15 meters in 24 mandals and between 5-10 meters in another 14 mandals. Situation in the urban areas also is no better [40(xi)].

It is in our religious practice to look to the skies and pray to the God of Rain (Varun) for his benevolence, thus induce him to bless us with good rains. A similar but a scientific method of inducing rains artificially from the skies, is found by the scientists which is called the 'Cloud Seeding Process'. This *Cloud Seeding Experiment* has been identified as one of the ways to mitigate this problem and has thus come to the rescue of the farmers of this district, to improve rainfall in this area [21]. Several progressive countries like USA[45], Australia, South Africa [6], China, Thailand, European countries, former states of USSR, Latin American countries, Arab countries, Indonesia, Israel [5],[13],[44] and Pakistan are getting highly benefited by employing the advanced cloud seeding technologies for over 40 years [13], [16], [29] & [34].

It is an established fact that, conducting Cloud Seeding Experiment is a costly affair [9] and hence, the State Government by itself took up the initiative in this endeavour and has been conducting many Cloud Seeding Experiments in Anantapur district for the last two years (vide, Appendix–III). These efforts of the Government could yield limited results (majority of the times 'nil' report) [40], not only because, during the, time of experimentation very thin clouds were available, but also because the selection of opportune time for conducting the Cloud Seeding Experiments was not well understood. Though, Studies have been going on through several agencies world over to get
successful results of the Cloud Seeding Experiments, in our country precious little is done so far. It is generally believed that, 'Thick formation of Clouds' is very essential besides many other factors, for Cloud Seeding Experiments to be successful. Without suitable cloud conditions, Cloud Seeding Experiments would not yield fruitful results. As large amounts of expenditure are involved in Cloud Seeding Experiments, identifying favourable climatic conditions, increases the chances for positive results and at the same time minimise the expenditure [22].

The above facts acted as natural motivation for us, to select this topic for Research and study the stochastic behaviour of rainfall of this district in order to predict the best appropriate time or Optimum Opportune Time (O.O.T) for conducting Cloud Seeding Experiments in this district, so that, the experiments will yield good rainfall which increases surface water, as well as, ground water resources. To meet this desired objective, the thesis concentrates on collection of data on rainfall of this district from 1st June 1996 – 31st Dec 2003 from three appropriately selected mandals viz., i) Amadagur, ii) Bommanahal and iii) Garldinne [48, 49].

This researcher earnestly believes that, the present endeavour will make a good beginning to assess the existing water resources, rainfall distribution, areawise as well as monthwise and to finally analyse the available data with the contemporary Statistical Techniques. With vagaries of the monsoon in general and in this area in particular, this study could give some quantitative aspects to clearly address this issue and to approach the Cloud Seeding Operations with a better foundation. It is firmly believed that, the statistical treatment of the
existing data will balance the inconsistencies in the rainfall distribution both laterally and in the time domain, and give some useful conclusions. This will help in identifying the **Optimum Opportune Time (O.O.T.)** and thus, planning the **future Cloud Seeding Operations during a more suitable time frame** to harness the unexploited and unlimited sky-water resources available in the atmosphere [61].

**Method of measuring rainfall in practice:**

Rainfall is usually measured by *'Rain Gauging Equipment'* located at mandal headquarters. Inherently, there are some limitations in this method, as the rain water collected in the rain gauging machine may not be representative and accurate to measure/project the rainfall of the entire mandal. Again, this is because of the fact that, rainfall is not uniformly spread throughout the mandal. At some places it may be more, and at other places it may be less. Hence, this measuring mechanism is not considered an appropriate one, to estimate the rainfall of entire mandal. In order to balance this inconsistency, it is felt that, a **concomitant variable** is to be identified to measure the rainfall, in addition to the **main variable**, which is the measured *rainfall* itself [[7], [38]].

**Identification of the concomitant variable:**

Dams are built across all flowing water bodies to harness the water resources for channelised utilization through a network of canals for catering to the needs of the people and that of agriculture. The levels of water reserves at Dam/Reservoirs are measured regularly and systematically recorded appropriately, in order to monitor its utilisation. Hence, **dam/reservoir levels**
can be considered as an appropriate indicator of the measure of rainfall, in the catchment areas of that dam/reservoir. Thus, in this thesis, in order to predict the rainfall of this district, we have considered various dam/reservoir levels of Anantapur and neighbouring areas, which are supplying water resources for drinking and agricultural purposes to this district [40].

There are two dams located in this district namely:

i) Mid-Pennar Reservoir (M.P.R. Dam) and

ii) Penna Ahobilam Balancing Reservoir (P.A.B.R. Dam).

In addition to these, water is supplied to Anantapur, Kurnool and Kadapa districts from Thunga-Badra Dam (T.B. Dam) located in the neighbouring Karnataka state through,

i) High Level Canal to Anantapur district (H.L.C)

ii) Kurnool and Kadapa Canal.

Hence, for Anantapur district, limited pre-determined water is supplied in a fixed period of time each year, from the High Level Canal system. These above water resources are catering to the needs of only a fraction of the population of this district. Majority of the people of this district, who live beyond the approach of the above canal systems, mainly depend on rainfall only. If appropriate remedial steps are not taken in the immediate future to improve rainfall of this district, soon it will be pushed to acute drought conditions, almost resulting in desert like situations. In this scenario, the Cloud Seeding Experiments, hold a great promise for this area of persistent draughts and looks like a ray of hope to the people of this district, and can prevent a major disaster [40].
Main Objective of this Thesis:

The success of the Cloud Seeding Experiments depends on selection of appropriate time and suitable cloud conditions during the Cloud Seeding Operations. Close understanding and analysis of the recorded rainfall in the immediate past and also the available water resources in the reservoirs located in this area will help to identify the appropriate time and suitable cloud conditions for these experiments. This fact has been emphasized by studies made by various countries that have already advanced further in this direction. The same has been highlighted by one of news items in the Hindu released on October 18–2006 [40(i)], which informs clearance of the experts panel to conduct Cloud Seeding Experiments in Andhra Pradesh for 27 days. Cloud Seeding Experiments conducted at an appropriate time will result in increased rainfall, which is essential for improving Surface water and Ground water resources so that the people of this district can lead a comfortable life. Hence, ‘Determination of Optimum Opportune Time (O.O.T) for conducting Cloud Seeding Experiments in this district’ is very essential and this most crucial concept should be borne in mind by the personnel/agencies conducting the Cloud Seeding Experiments [55].

The main objective of this Thesis is to determine the O.O.T for conducting Cloud Seeding Experiments in this district, based on the data collected on rainfall and water levels in the dams located in this district, using various statistical tools, techniques and models. These types of efforts are highly useful and reliable to judiciously plan and utilize public funds in an appropriate and fruitful manner for the benefit of the people of Ananatpur district.
Now, in the following sub-sections, we proceed to elaborate on some definitions, preliminary concepts and the methodology of conducting *Cloud Seeding Experiments*.

### 1.1. a) Definitions, Basic concepts related to Cloud Seeding Experiments:

**Cloud Seeding:**

*Cloud Seeding* is also commonly known as Weather modification, Cloud modification, Atmospheric resource management and Precipitation management [[1], [8]].

**Need for Cloud Seeding** [[15], [29]]:

The key role played by a good water supply as an engine of economic growth and as a yard stick of public welfare and national prosperity has been well recognized by the intellectuals of the developed countries like USA, who aptly named water as the 'Blue Gold'. Hence, the advanced countries are constantly upgrading their water resources by harnessing not only all the ground and surface waters but also by tapping a renewable, virtually unlimited and unexploited sky-water resource available in the atmosphere in the form of innumerable clouds.

Unfortunately, the vagaries of Indian monsoons very frequently cause floods in the East and North Eastern states, while the Western and Southern states face recurring droughts which cause water scarcities that are adversely affecting the health of the human and animal populations, food production, hydro-power generation and industrial growth. To mitigate the dire
consequences of drought and to face the immediate challenges of water crisis, cloud seeding experiments seem to hold a ray of hope for various purposes like:

i). Increase of annual rainfall for drinking and agricultural purposes,

ii). Dispersal of fog in airports and city roads

iii). Increase of hydro-power generation at the cheapest cost

iv). Suppression of hail storms to reduce damage to life, crops and properties

v). Mitigation of devastating impacts of recurring droughts

vi). Mitigation of damaging impacts of summer temperatures

vii). Increase of annual rainfall for improving the forests, wildlife and the environment [[15], [20] & [21]].

Several Indian states interested in promoting economic growth, agriculture development and public welfare are eager to learn from the successful experiences of other countries like China and USA and adopt those technologies by making necessary modifications to suit the local meteorological, topographical, geographical and other environmental conditions [33].

Principles of Cloud Seeding:

Clouds are made up of millions and millions of water droplets or tiny ice particles or both, which form around microscopic particles of dust, smoke, soil, salt crystal and other chemical aerosols, bacteria and spores that are always present in the atmosphere. These particles are classified as ‘condensation nuclei’ (CCN) [2] on which water vapour condenses to form cloud droplets and a few of them are classified as ‘Ice Nuclei’ (IN) [[12], [30]] on which condensed water freezes or ice crystals form directly from water vapour. In the normal
atmosphere, there is an abundance of condensation nuclei while there is a scarcity of ice nuclei. The types of nuclei and their sizes and concentrations present in the air play a significant role in determining the efficiency with which a cloud system precipitates. Generally tonnes and tonnes of water flow, as rivers of moisture, in the skies over many countries and from these rivers in the sky either little precipitation or no precipitation falls on the ground, because of absence of certain required conditions [17]. Among such important conditions for both initiation of precipitation and the amount of precipitation from a cloud system are:

i). Horizontal and vertical dimensions of cloud

ii). Lifetime of the cloud and

iii). Sizes and concentrations of cloud droplets and ice crystals [31, 36 & 37].

Under proper conditions one or more of these above 3 factors can be favorably modified by seeding the clouds with appropriate nuclei, mostly by using either common salt or silver iodide particles [[30], [32]].

**Warm and Cold Cloud Seeding:**

Precipitation forms in clouds by two mechanisms namely ‘warm rain’ and the ‘cold rain’ processes. The term warm rain was coined by the scientists who found that the rain in tropical countries often fell from clouds whose temperature throughout the clouds was warmer than the freezing level of 0°C or 32°F. Rain occurs in these clouds when larger droplets collide with the smaller cloud droplets and absorb them in a process known as ‘coalescence’. The cold rain occurs in clouds whose temperature in all or part is colder than the freezing level of 0°C or 32°F. The regions of the cloud below the freezing level
are super cooled and contain both water droplets and ice crystals and sometimes only the former. The ice crystals which form in the super cooled regions of the cloud grow very rapidly by means of drawing the moisture from the surrounding cloud droplets and this growth continues until their weight overcomes the gravity forces and causes them to fall to the ground. While falling from cloud these ice crystals coalesce with other smaller droplets and fall from the cloud as snow or rain \([1], [10] \& [28]\).

The atmospheric nuclei that play a key role in cloud formation exert a strong influence on the efficiency with which the warm and cold rain processes operate. For example, the giant condensation nuclei are prevalent in the oceanic atmosphere that allows for larger cloud droplets to form and the coalescence process to initiate rain within the life time of the cloud. But the continental regions are characterized by much smaller and number of condensation Nuclei. Hence medium size clouds formed over the continental areas generally dissipate before the coalescence mechanism has had a chance to initiate rain. Similarly, many regions have a shortage of Ice Nuclei which reduces the efficiency of cold rain process [10].

**Reasons for Injecting Chemicals into Clouds:**

Under favourable natural environmental conditions clouds can be stimulated to grow larger and also to last longer. The injection of silver iodide particles into the super cooled part of the cloud makes the cloud droplets freeze into ice crystals. This conversion process gets multiplied millions and millions of times within the cloud and releases a large amount of heat, known as the
'latent heat of fusion'. This phenomenon makes the cloud more buoyant and makes it grow larger in size and thereby makes the cloud process more efficient and longer than that of a naturally formed cloud [35].

**Clouds and their characteristics:**

**Cloud formation:**

A cloud is nothing but a visible conglomeration of very small particles of water or ice or of both in the atmosphere. The moisture that goes into the atmosphere is due to heating up of the water from the surface of lakes, streams and oceans by the Sun and it exists in the form of invisible vapour [41].

**Properties of clouds:**

If continued moisture supply is available more and more condensation will occur, resulting in the formation of fog or mist. When this fog is lifted above the ground, a sheet like cloud known as **stratus cloud will be formed**. If this condensation occurs above the ground level after some moist air is forced vertically upwards as updrafts, then small heaps of clouds form and these clouds may not have a growing tendency and hence may dissipate without giving rain. However when large areas in the sky are covered with streets of puffy **cumulus (heap)** clouds, the larger ones among them with intense updrafts or vertical upward motion of air at a rate of 5 to 10 meters per second, the clouds may grow by as much as 6km in thickness, known as **towering cumulus clouds**. These clouds contain water droplets only in super cooled state. Additional growth of such clouds may result in the freezing of water droplets. When the top surface of these clouds grow upto about 16km and assume anvil
shapes indicating outflow of air from the top of the clouds, such clouds are known as **cumulo-nimbus (violent rain) clouds** [8], [13], [30], [39] & [46].

The shape of the cloud and its appearance are determined by the nature, number and size of nuclei and droplets and the weather characteristics. The clouds are always in a continuous evolution process and hence display many varieties of forms. Slow and prolonged ascent of air in a low pressure area or irregular stirring motion produces a layer type of clouds known as **stratiform clouds**. Convective currents leading to violent ascent of air masses due to insulation caused by heat from the sun create heaps of clouds known as **cumuliform clouds**. All the clouds can be divided into four classes based on the thickness and height of the cloud base as shown below:

**Table (1.1.1): Classification of Clouds**

<table>
<thead>
<tr>
<th>Class of Clouds</th>
<th>Name of Clouds</th>
<th>Height of base</th>
<th>Temp. at base</th>
<th>Main constituent</th>
<th>Precipitation</th>
<th>Appearance</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>High Clouds</strong></td>
<td>Cirrus (Ci)</td>
<td>5 - 14 km</td>
<td>-20°C to -60°C</td>
<td>Ice</td>
<td>No</td>
<td>White Philaments, fibrous</td>
</tr>
<tr>
<td></td>
<td>Cirrocumulus (Cc)</td>
<td>5 - 14 km</td>
<td>-20°C to -60°C</td>
<td>Ice</td>
<td>No</td>
<td>White patches or layers</td>
</tr>
<tr>
<td></td>
<td>Cirrostratus (Cs)</td>
<td>5 - 14 km</td>
<td>-20°C to -60°C</td>
<td>Ice</td>
<td>No</td>
<td>Translucent Cloud veil</td>
</tr>
<tr>
<td><strong>Medium Clouds</strong></td>
<td>Altocumulus (Ac)</td>
<td>2-7km</td>
<td>+10°C to -30°C</td>
<td>Water with some ice</td>
<td>Occasional</td>
<td>White grey Layers</td>
</tr>
<tr>
<td></td>
<td>Altostratus (As)</td>
<td>2-7km</td>
<td>+10°C to -30°C</td>
<td>Water with some ice</td>
<td>Occasional</td>
<td>Grey sheets with ice in row</td>
</tr>
<tr>
<td><strong>Low Clouds</strong></td>
<td>Nimbostratus (Ns)</td>
<td>1-3km</td>
<td>+10°C to -15°C</td>
<td>Water and ice</td>
<td>Good rain</td>
<td>Dark grey layers</td>
</tr>
<tr>
<td></td>
<td>Stratocumulus (Sc)</td>
<td>0.5-2km</td>
<td>+15°C to -5°C</td>
<td>Water</td>
<td>Little rain</td>
<td>Grey patchy sheets</td>
</tr>
<tr>
<td></td>
<td>Stratus (St)</td>
<td>0.5km</td>
<td>+20°C to -5°C</td>
<td>Water</td>
<td>Drizzle</td>
<td>Grey uniform layer</td>
</tr>
<tr>
<td><strong>Clouds with vertical growth</strong></td>
<td>Towering cumulus or large cumulus</td>
<td>0.5-2km</td>
<td>+15°C to -5°C</td>
<td>Water</td>
<td>Small showers</td>
<td>Detached clouds grow vertically</td>
</tr>
<tr>
<td></td>
<td>Cumulonimbus (Cb) or Thunderstorm</td>
<td>0.5-2km</td>
<td>+15°C to -5°C</td>
<td>Water</td>
<td>Heavy rain</td>
<td>Huge towering clouds</td>
</tr>
</tbody>
</table>
Clouds play an important role in weather forecasting. Feather like **Cirrus (curls of hair) clouds** in the middle and low latitudes scattered in the sky during summer indicate clearing of the weather [[3], [9], [16], [46] & [62]].

**Rain-clouds:**

When **cumulus clouds** gradually develop into cumulo-nimbus clouds on hot and humid days, they indicate an approaching thunder storm and hence cloud seeding has to be planned in advance to prevent the development of hail storms that cause severe damage to crops and properties. When the **stratocumulus clouds** cover the morning sky and the wind appears gusty, clouds will develop during the day time. Cloudiness and rainfall are interrelated in the tropical equatorial belt where there exists substantial cloud cover with maximum amount of rainfall. The clouds are localised and have great vertical development. Due to the descending air currents in the subtropical belt, rainfall and cloud cover will be less and deserts are located in such regions. In the equatorial belt, cloudiness does not vary much from month to month. But the cloud maximum occurs between 10° and 20° North and South latitudes during summer months with substantial rainfall. Daily variations in cloudiness are based on the kind of clouds in the sky. The maximum cumulus or cumuliform clouds occur during the early and middle afternoon periods while stratus and **stratiform clouds** appear to the maximum extent during the morning periods. While **altostratus clouds** give light rain, **towering cumulus clouds** provide short spells of showers. While Nimbostratus clouds give moderate continuous rain, **cumulonimbus clouds** give heavy rain, often accompanied by hail stones [[58], [61]].
Several other kinds of clouds do not provide rain. Stratus clouds are in the lowest level in the sky. When there is a depression within 200 to 300 kms from the sea coast stratus clouds in the shape of cotton pieces moving fast, touching the tall trees are found along the seacoast. This cloud is always thin and has a stratiform or horizontal top. If the depression close by is intensifying further, hundreds of these stratus clouds are seen moving like an invading army [60].

**Cumulus clouds** are 5000 ft thick and are generated by the summer heat when some moisture is present in the atmosphere. They disappear by night when the convection currents subside. They have rugged or cauliflower shaped tops and are known as fair weather cumulus clouds. During summer when there is moist air and good convection cumulus clouds grow vertically and reach the freezing levels around 6km. Super cooled water often triggers sharp rainfall that lasts upto half an hour. If there is stable upper air layer, the top of the cumulus cloud gets flattened and the cloud becomes a strato-cumulus cloud. Small strato-cumulus clouds do not shed any rain. Alto-cumulus clouds when closely packed look like a flock of sheep and they do not give any rain. If the sky is covered with alto cumulus clouds for 2 to 3 days it indicates that a low pressure system within 300 km is developing into a depression or a cyclone. Cyclonic storm winds associated with a depression cause a large scale convergence for about 300 kms around the depression at the lower levels of the atmosphere. **Altostratus clouds** then form with a base around 8000 ft and a thickness of 5000 ft to 10,000ft and its spread cloud be around 300 kms around the depression. By its gradual growth and thickness, it provides
corresponding increase in rainfall reaching up to heavy downpour which continues until the depression disappears. The rain may last for 2 to 3 days.

The high clouds that form above 7 km in the sky are known as **cirrus clouds** that are very thin and contain ice-crystals. These clouds do not give any rain. The thunder cloud, the giant king of the sky, is known as **cumulo-nimbus cloud** which is classified as low cloud with its base around 1.5 km and its top reaching 12 km in the sky. They provide rain for about half an hour in dry summer and they dominate the sky during monsoon periods for 2 to 3 hours and produce heavy rains. The **Nimbo-stratus cloud**, a very dark stratiform cloud gives moderate rainfall for a short time. Among the clouds, the most important ones that provide rainfall are the tall cumulus, the large Strato cumulus, the thick Alto-stratus and the dark Cumulo-nimbus. Since the human eye can observe the horizon up to 40 kms, it is necessary to cultivate sky reading to predict the rain which is influenced both by the vertical extent as well as the horizontal dimensions of the clouds.

**Suitable clouds for Cloud Seeding Experiments** [46]:

All the clouds are not suitable for the **Cloud Seeding** operations. The cloud seeding operations themselves cannot be done if there are no clouds in the atmosphere and hence, operators choose Cumulo-Nimbus clouds which are convective in nature with a great deal of vertical mixing. Nimbo stratus clouds and cumulus clouds which are dark grey are also suitable. The cloud must be deep enough with temperatures within a suitable range for seeding. There should be significant levels of super cooled liquid water present in the cloud.
The wind also must be below a specified value. Since it takes about half an hour for the artificially injected chemical crystals to grow into raindrops, seeding line must be 30 minutes upwind of the selected target area boundary. If the speed of the wind is 40 knots the seeding gets done 20 nautical miles or 37km upwind of the target area boundary. **If all the criteria are met before launching the operations, cloud seeding becomes successful in producing substantial additional rainfall** [33].

Cloud seeding operations are conducted during the rainy season in suitable places and the operations take place when it is safe to fly and suitable weather conditions are present, 24 hours a day and 7 days a week. The base of the cloud must be within 1.5km to 2km from the ground. A reasonable degree of local updrafts promote the augmentation of rainfall.

**Cloud Seeding experiments in India:**

Indian Institute of Tropical Meteorology, Pune has conducted so many cloud seeding experiments throughout India, to obtain Artificial rains. These cloud seeding experiments were conducted along with the flow of clouds during the South-West monsoons and could succeed in increasing of rainfall by 25% in addition to the actual rainfall. In this direction Karnataka State also conducted the cloud seeding experiments, since 1975 and succeeded in getting additional rainfall.

At present, Karnataka State Government is conducting these cloud seeding experiments under the name of ‘**Project-Varuna**’. Under this project they got 73% success in getting additional rainfall. Impressed by the success of
the cloud seeding operations, obtained in Karnataka, neighboring States like, Andhra Pradesh and Maharashtra have also decided to go on, to do similar type of exercises to get additional rainfall by conducting cloud seeding experiments. As per the available sources, Maharashtra Government has signed an agreement with *Weather Modification Inc.* and its Indian representative *Agni Aviations* for a 90-day exercise to obtain artificial rains.

Andhra Pradesh Government also started conducting these experiments of and on, since 1993, but these experiments could not give good results. Again in 2004, the Andhra Pradesh Government has started cloud seeding experiments in the name of *'Indira Megha Madhanam'* and signed an agreement with *Weather Modification Inc (WMI)* and its Indian Representative *Agni Aviations* to obtain artificial rain on similar lines. These experiments were spread over areas covering 10 rain-shadow districts of the state, for a period of 126 days in 5 months. The districts covered under the scheme are *Anantapur, Kadapa, Kurnool, Chittoor, Nellore, Guntur, Nalgonda, Mahaboobnagar, Ranga Reddy* and *Prakasam*. For the 4 Rayalaseema districts, a radar station was established at *Anantapur town* itself with an ‘attached airport’ facility at Puttaparthi for the movement of the aircraft to Bangalore and Kadapa. For the rest of the areas, which are located in Telangana region, the radar was set up at Ibrahimpatnam. In these experiments silver chloride, silver bromide / iodide are used as a seed material to precipitate the clouds.

An appraisal of the Cloud Seeding Operations carried out during 2006 is believed to have been successful and has enhanced the rainfall over the areas covered in 10 districts. Out of a total of 12,160 mm rainfall claimed as received...
by these areas in 10 districts, 1,825 mm has been attributed to the Cloud Seeding Operations as informed by the Centre for Atmospheric Sciences & Weather Modification Technologies (CASWMT) of JNTU which is overseeing these operations [40(x)]. Also, it has been observed by this agency that 42 out of the 50 rainwater samples collected from the areas were found to contain calcium and potassium at higher levels indicative of result of cloud seeding operations [40(i)].

**Advantages:** The advantages of this cloud seeding technology are as follows:

- It contributes to the increase in freshwater resources to meet the growing water demands. Efficiency can be increased through relevant research, experiments and a better understanding of precipitation and cloud systems.

- Cloud seeding operations under certain conditions have produced positive results, using either stratus clouds formed by the collision of moist air masses over high mountain regions or convective clouds.

- It succeeded in the dissipation of fogs and low stratus clouds that cause obstacles to aviation traffic at airports and automobile traffic in cities.

- It improves the productivity in rain-fed agricultural fields, either by increasing the rainfall amounts or sometimes by controlling the spatial and temporal distribution of rainfall [39].
Disadvantages: The disadvantages of this technology are as follows:

- The success of this technology under drought conditions is very limited, because of the absence of seedable clouds during such periods and hence cloud seeding must be done when favourable clouds are available to store the water for later utilization.

- Cloud seeding operations require advanced and costly equipment, and also qualified professional experts like good pilots and skilled meteorologists [39].

Future Development of the Cloud Seeding Technology [[27], [42] & [43]]:

The most important aspect of cloud seeding technology is its development as a tool to solve the emerging problems of water scarcity in arid and semi-arid regions of the world [59]. Some of the methods to be adopted to develop and improve the technology are as follows:

- Implementation of joint cloud seeding projects between neighboring states and countries, with the assistance of international organizations in this field and utilizing the results of successful experiments from other countries.

- Conducting cloud surveys for determination of their precipitation potential and analysis of cloud-seeding methods and operations.

- Collection of data on cloud characteristics and cloud seeding operations for publication and distribution among the institutions and countries interested in promoting operations and research in cloud-seeding.
• Promoting the participation of beneficiaries like farmers, and industrialists in cloud seeding to achieve positive results for augmenting water availability required for agriculture and industries. Since cloud seeding results are tangible and multi-faceted, attention must be given not only to meteorology and water, but also to the consequential ecological, hydrological, social and economic problems and the project proponents must inform the decision-makers in the Government and the general public of the current weather modification technologies and their impacts on regional and national development.

• Indepth study and research on Cloud Seeding Experiments is necessary to determine **Optimum Opportune Time (O.O.T.)**, based on latest trends/data collected, which is to be analyzed statistically [61].

• Suitable stochastic models are to be proposed to predict the behaviour of Clouds and rainfall so that one can predict O.O.T. for conducting Cloud Seeding Experiments [61].

• In this direction, one can search for a **concomitant variable**, through which the behaviour of the **main variable** namely **rainfall** can be predicted in any area of our interest.

The present thesis concentrates on the above points namely, identifying concomitant variable and predicting the behaviour of the rainfall through which O.O.T. for conducting Cloud Seeding Experiments are determined for Anantapur district. These points are explained elaborately in the following section.
1.2. Role of Rainfall, Dam/Reservoir levels in determination of O.O.T. for conducting Cloud Seeding Experiments:

In the previous section, we have discussed the importance of Cloud Seeding Experiments, some definitions and basic ideas. It is important to note that thick formations of clouds are essential for the successful conduct of Cloud Seeding Experiments. Rainfall is the basic fruit of Cloud Seeding Experiments, which is most essential for improving agricultural returns. Rainfall is a stochastic variable that varies from time to time in a very erratic manner. To study the stochastic behaviour of rainfall many approaches came into existence. Still the determination of rainfall at a particular place cannot be measured exactly through the measuring mechanism of rainfall. It is already mentioned in the last section that fixing the rain gauging machine at one place and collecting the rain water through that machine and measuring the rainfall may not be accurate one, to estimate the rainfall of entire area, because rainfall is not uniformly spread throughout the area. Hence, to measure the main variable rainfall (X), we take the help of another concomitant variable dam/reservoir levels (Y) which mainly depends on the rainfall. Further, records about the dam/reservoir levels are available from the records in a perfect and accurate manner. Records are maintained about the dam/reservoir levels measured in ‘Million Cubic Feet’ on daily basis. This data is more appropriate to measure the rainfall of the catchment area of the dam concerned. Thus, the data related to various dam/reservoir levels located at Anantapur, forms the main basis of the analysis of the thesis. Using dam/reservoir levels, rainfall is predicted and ‘most probable rainy months’ are determined. Most probable rainy months are appropriate for conducting Cloud Seeding Experiments.

22
The data about the dam/reservoir levels is collected from 'The Circle Office of Thunga Bhadra Project and High Level Canal' located at Anantapur, for the following 4 Reservoirs:

1) Thunga Bhadra Dam (T.B.Dam)
2) Mid Pennar Reservoir (M.P.R.Dam)
3) Mailavaram Reservoir (MLV.Dam) and
4) Penna Ahobilam Balancing Reservoir (P.A.B.R.Dam)

The data is collected for the period from '1st June 1996 – 31st May 2006', from the records maintained by the offices mentioned above.

In order to measure the relation between dam/reservoir levels and the rainfall of their respective catchment areas, rainfall data measured accurately in 'Mille Meters' is collected for the period form '1st June 1996 – 31st Dec 2003'. The above rainfall data is collected from the records maintained by Mandal Headquarters of Mandal Revenue Offices located at:

1) Amadagur
2) Bommanahal and
3) Garladinne

Above collected data on rainfall (X) and dam/reservoir levels (Y) during the period mentioned above, forms the main basis for Statistical Analysis in this thesis. The original data collected on rainfall (X) and dam/reservoir levels (Y) are provided in a CD form attached to this thesis.
1.3. Objectives of the Thesis:

It is already mentioned in Section-I that the basic objective of the thesis is to predict **O.O.T. for conducting Cloud Seeding Experiments** through various sophisticated statistical techniques available in literature. Thus, the objectives of this thesis are elaborated in the following paragraphs:

1. To calculate the relation between rainfall and dam/reservoir levels and to predict rainfall using **Linear Regression Model** and to predict O.O.T. for conducting Cloud Seeding Experiments [60].

2. To propose a **Markov Chain Model** for dam/reservoir levels, to calculate **Transition Probabilities** and **Steady State Solutions** for the determination of stochastic nature of rainfall and to predict Optimum Opportune Time for conducting Cloud Seeding Experiments [24].

3. To predict the rainfall through **Simulation Models** using Steady-state solutions and to predict an **appropriate time** for conducting Cloud Seeding Experiments through these proposed simulation models [14, 60].

4. To test the closeness between **Simulation Models** proposed and **Actual (observed) Models** in order to determine the rainfall [14].
1.4. Tools and Techniques used:

Keeping the above objectives in view, different statistical models have been applied as listed below:

1. ANOVA – One – Way Classification
2. Linear Correlation and Regression Models
3. Markov Chain Models and

A brief discussion about these models, tools and techniques is given below:

1.4 a) Analysis of Variance (ANOVA) [[51], [52]]:

Definition:

According to R.A. Fisher, Analysis of Variance is the separation of variance due to one group of causes from the variance due to another group of causes.

The total variation in any set of numerical data due to number of causes or the variation in the quality control may be classified into two causes.

They are:

i) Assignable causes

ii) Chance causes

In the ANOVA technique, the variation due to assignable causes can be estimated and compared with the variation due to chance causes. There are two types of ANOVA.

They are:

i) One-Way-Classification

ii) Two-Way-Classification
Analysis of One-Way-Classification:

Let the population N observations be divided into k classes of sizes, \( n_1, n_2, n_3, \ldots, n_k \). The classes are also termed as 'Treatments'. Let \( x_{ij} \), where \( i = 1, 2, \ldots, n_i \) is the \( j \)th value in the \( i \)th class, which are expressed in the following table:

<table>
<thead>
<tr>
<th>CLASSES</th>
<th>TOTAL</th>
</tr>
</thead>
<tbody>
<tr>
<td>( X_{11}, X_{12}, X_{13}, \ldots, X_{1n_1} )</td>
<td>( T_1 )</td>
</tr>
<tr>
<td>( X_{21}, X_{22}, X_{23}, \ldots, X_{2n_2} )</td>
<td>( T_2 )</td>
</tr>
<tr>
<td>( X_{31}, X_{32}, X_{33}, \ldots, X_{3n_3} )</td>
<td>( T_3 )</td>
</tr>
<tr>
<td>( \ldots )</td>
<td>( \ldots )</td>
</tr>
<tr>
<td>( X_{k1}, X_{k2}, X_{k3}, \ldots, X_{kn_k} )</td>
<td>( T_k )</td>
</tr>
</tbody>
</table>

The total variation in the observations \( X_{ij} \) can split into two components. They are, the variation between the classes, which is commonly known as treatments, and the variance due to treatments which is due to chance causes. The main object of analysis of variance is to examine the significant difference between the class means. Let the class means of population be \( \mu_1, \mu_2, \mu_3, \ldots, \mu_k \). Then the hypothesis is

\[
H_0: \mu_1 = \mu_2 = \mu_3 = \ldots = \mu_k = \mu
\]

MATHEMATICAL MODEL:

The linear mode is

\[
X_{ij} = \mu_j + e_{ij}
\]

\[
= \mu + (\mu_j - \mu) + e_{ij}
\]

\[
X_{ij} = \mu + a_i + e_{ij} \quad \text{----------------------------- [1.4.1]}
\]
Where $\mu = \text{general mean effect}$

$\alpha_i = \text{Effect due to } i^{th} \text{ treatment}$

$\mu_i = \text{mean of the } i^{th} \text{ class population}$

$e_{ij} = \text{Error effect due to chance}$

**Assumptions:**

1) All the observations are independent.

2) Different effects are additive in nature.

3) $e_{ij}$'s are i.i.d $N(0,\sigma^2)$

**Analysis:**

$$\bar{X}_{i.} = \text{mean of the } i^{th} \text{ class} = \frac{\sum_{j=1}^{n_i} X_{ij}}{n_i}$$

$$= \frac{T_i}{n_i} \quad \text{[1.4.2]}$$

$$\bar{X}_{..} = \text{overall mean} = \frac{\sum_{i=1}^{k} \sum_{j=1}^{n_i} X_{ij}}{N}$$

$$= \frac{G}{N} \quad \text{[1.4.3]}$$

Where $N = \sum_{i=1}^{k} n_i$, $G = \text{grand total} = \sum T_i$

Total sum of squares (TSS) $= \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X})^2$

$$= \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{i.})^2 + \sum_{i=1}^{k} \sum_{j=1}^{n_i} (X_{ij} - \bar{X}_{..})^2$$
TSS = Sum of squares of error + Sum of squares of treatments

\[ TSS = SSE + SST \]  \[ \text{[1.4.4]} \]

\[ ESS = TSS - SST \]  \[ \text{[1.4.5]} \]

**Degrees of freedom:**

Degrees of freedom for total = \( N - 1 \)

Degrees of freedom for treatment = \( k - 1 \)

Degrees of freedom for error = \((N-1)-(k-1) = N-k\)

**Mean Sum of Squares (MSS):**

MSS are obtained by dividing the sum of squares by their respective degree of freedom.

\[ \text{MSS for treatments} \quad (M^2_t) = \frac{SST}{K-1} \]  \[ \text{[1.4.6]} \]

\[ \text{MSS for error} \quad (M^2_e) = \frac{SSE}{K-k} \]  \[ \text{[1.4.7]} \]

\[ F - \text{Ratio} : \quad F = \frac{M^2_t}{M^2_e} \sim F \ (k-1, \ N-k) \]  \[ \text{[1.4.8]} \]

**Table (1.4.2): ANOVA Table of One-Way-Classification:**

<table>
<thead>
<tr>
<th>Source of variation</th>
<th>Degrees of freedom</th>
<th>Sum of squares</th>
<th>Mean sum of squares</th>
<th>F - calculated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatments</td>
<td>( K - 1 )</td>
<td>SST</td>
<td>( M^2_t = V_1 )</td>
<td>( F = \frac{M^2_t}{M^2_e} \sim F \ (k-1, \ N-k) )</td>
</tr>
<tr>
<td>Error</td>
<td>( N - k )</td>
<td>SSE</td>
<td>( M^2_e = V_2 )</td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>( N - 1 )</td>
<td>TSS</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

At \( \alpha \% \) l.o.s,

If \( F - \text{calculated} \leq F \ (k-1, \ N-k) \Rightarrow \text{we accept } H_0 \)

Otherwise we reject \( H_0 \).
FURTHER STATISTICAL ANALYSIS [53]:

If \( H_0 \) is not rejected in ANOVA, we conclude that there is no significant difference between the treatment means and analysis stops there because all the treatments are equally effective and one can use any one of the treatments. But, if \( H_0 \) is rejected concluding that there exists significant difference among the treatment means, it will not stop the analysis, because one may be interested to know the best treatment among the given treatments. Perhaps we are interested to know the treatments, which are significantly different, which are not significantly different among the given treatments. To do this, we have to analyze further to subgroup the treatments, such that treatments within a subgroup have no significant difference and treatments between the subgroups have significant difference. The sub grouping is done by the following tests:

1) Critical Difference Test (CD – Test) / Least Significant Test (LSD Test)

2) Student – Neymann – Keul’s Test (SNK – Test)

3) Duncan’s Multiple Range Test (DMR – Test)

To apply these tests, first subclass means are to be calculated and are to be arranged in an order.

Let \( \bar{y}_1, \bar{y}_2, \bar{y}_3, \ldots, \bar{y}_k \) are treatment means of \( T_1, T_2, T_3, \ldots, T_k \).

Now, arrange these means in an ascending or descending order as \( \bar{y}_{(1)}, \bar{y}_{(2)}, \ldots, \bar{y}_{(k)} \)

Critical Difference (CD) Test or Least Significant Difference (LSD) Test:

Let \( \bar{y}_l \) be the \( i^{th} \) subclass mean,

\[
V (\bar{y}_l) = V (\frac{\bar{y}_l}{r})
\]
\[
V(\bar{y}, -\bar{y}; \frac{2\sigma^2}{r}) = \frac{1}{r^2} \sum V(y_i)
\]
\[
= \frac{1}{r^2} \sum \sigma^2 = \frac{\sigma^2}{r}
\]

Similarly \(V(\bar{y}_i, -\bar{y}_j) = \frac{2\sigma^2}{r}\) \[1.4.9\]

\[
SE(\bar{y}_i, -\bar{y}_j) = \frac{\sigma}{\sqrt{r}} \quad \text{-------------------------} \quad [1.4.10]
\]

If the number of units are unequal then
\[
V(\bar{y}_i, -\bar{y}_j) = \sigma^2\left(\frac{1}{r_1} + \frac{1}{r_2}\right) \quad \text{-------------------------} \quad [1.4.11]
\]

\[
SE(\bar{y}_i, -\bar{y}_j) = \sigma \sqrt{\frac{1}{r_1} + \frac{1}{r_2}} \quad \text{-------------------------} \quad [1.4.12]
\]

To test the hypothesis \(H_0: \bar{y}_i = \bar{y}_j\), we are using t-test

The test statistic,
\[
t = \frac{\bar{Y}_i - \bar{Y}_j}{SE(\bar{Y}_i, -\bar{Y}_j)} \sim t_{K(r-1)} \quad \text{-------------------------} \quad [1.4.13]
\]

\[\therefore\] If \(t_{\text{cal}} < t_{\text{table}}\) \(\Rightarrow\) we accept \(H_0\),

Otherwise reject \(H_0\).

Define \(CD = D_\alpha = t_{K(r-1)} \left[SE(\bar{y}_i, -\bar{y}_j)\right]\)

\[
= t_{\text{table}} \ast \frac{\sigma}{\sqrt{r}} \quad \text{-------------------------} \quad [1.4.14]
\]

In ANOVA table, we get \(\sigma = \sqrt{\frac{2}{r}}\) \[1.4.15\]
Now calculate the difference between $\bar{y}_{(i)}$ and $\bar{y}_{(k)}$

i.e., $\bar{y}_{(i)} - \bar{y}_{(k)}$

If $\bar{y}_{(i)} - \bar{y}_{(k)} < D_a$, we conclude that there is no significant difference between 1 to k treatments.

If $\bar{y}_{(i)} - \bar{y}_{(k)} > D_a$, then calculate:

Case (i): $\bar{y}_{(k-1)} - \bar{y}_{(1)} \leq D_a$, then we will have $\{t_1, t_2, \ldots, t_{k-1}, t_k\}$

Case (ii): $\bar{y}_{(i)} - \bar{y}_{(2)} \leq D_a$, then we get $t_1, (t_2, \ldots, t_{k-1}, t_k)$

Repeat this process and conclude by sub grouping all the treatments.

This procedure basically depends upon calculating the difference between the means which are arranged in order and comparing this difference with $D_a$, where $D_a$ is called critical difference at $\alpha$ % l.o.s. sometimes $D_a$ is known as least significant difference. This test is also known as Least Significant Difference (LSD Test). In this thesis we have applied CD/LSD test and this test alone is explained here.

1.4. b) Linear Correlation and Regression Model [[51] & [52]]:

Correlation:

Correlation is a statistical device which helps us in analyzing the co-variation of two or more variables. For example, there exists some relationship between the age of husband and the age of wife, price of the commodity and the amount demanded, etc.
**Definition:**

When the relationship is of a quantitative nature, the appropriate statistical tool for discovering and measuring the relationship and expressing it in brief formula is known as correlation.

----- Croxton & Cowden

The degree of relationship between the variables under consideration is measured through the correlation analysis. The correlation analysis refers to the techniques used in measuring the closeness of the relationship between the variables. The measure of correlation is called the **coefficient of correlation**.

**Karl Pearson’s coefficient of Correlation:**

Of the several mathematical methods of measuring correlation, the Karl Pearson’s method, popularly known as Pearson’s coefficient of correlation, is most widely used in practice. Correlation coefficient between two random variables X and Y, usually denoted by r(X, Y) or simply r_{XY}, is a numerical measure of linear relationship between them and is defined as

\[
 r(X, Y) = \frac{Cov(X, Y)}{\sigma_x \sigma_y} \quad \text{[1.4.16]}
\]

if \((x_i, y_i)\) \(i = 1, 2, 3, \ldots n\) is the bivariate distribution, then

\[
 Cov (X, Y) = \frac{1}{n} \sum (x_i - \bar{x})(y_i - \bar{y})
\]

\[
 \sigma_x^2 = \frac{1}{n} \sum (x_i - \bar{x})^2, \quad \bar{x} = \frac{\sum x_i}{n}
\]

\[
 \sigma_y^2 = \frac{1}{n} \sum (y_i - \bar{y})^2, \quad \bar{y} = \frac{\sum y_i}{n}
\]
Assumptions of Karl Pearson's Correlation Coefficient:

Karl Pearson's coefficient is based on the following assumptions:

- There is linear relationship between the variables.
- The two variables under study are affected by a large number of independent causes so as to form a normal distribution.
- There is a cause and effect relationship between the forces affecting the distribution of the items in the two series. Such a relationship is not formed between the variables, i.e., if the variables are independent.

Limitations:

- The correlation coefficient always assumes linear relationship regardless of the fact whether that assumption is correct or not.
- Great care must be exercised in interpreting the value of this coefficient as very often the coefficient is misinterpreted.
- The value of the coefficient is unduly affected by the extreme items.
- As compared with other methods this method takes more time to compute the value of correlation coefficient.

Interpreting Coefficient of Correlation:

The following general rules are given which would help in interpreting the value of ‘r’:

- When \( r = +1 \), it means there is perfect positive relationship between the variables.
• When \( r = -1 \), it means there is perfect negative relationship between the variables.

• When \( r = 0 \), it means that there is no relationship between the variables i.e., the variables are uncorrelated.

• The closer \( r \) is to +1 or -1, the closer the relationship between the variables and the closer \( r \) is to 0, the less close the relationship. The full interpretation of \( r \) depends upon circumstances, one of which is the size of the sample. All that can really be said that when estimating the value of one variable from the value of another, the higher the value of \( r \) the better the estimates.

• The closeness of the relationship is not proportional to \( r \). If the value of \( r \) is 0.8 it does not indicate a relationship twice as close as one of 0.4. It is, in fact, very much closer.

**Properties of the Coefficient of Correlation:**

The following are the important properties of the correlation coefficient \( r \):

1. The coefficient of correlation lies in between -1 and +1.

2. The coefficient of correlation is independent of change of scale and origin of the variables.

3. The coefficient of correlation is the geometric mean of two regression coefficients.

4. The degree of relationship between the two variables is symmetric.
**t-test for testing significance of an observed sample correlation coefficient:**

If 'r' is the observed correlation coefficient in a sample of 'n' pairs of observations from a bivariate normal population, then **Prof. Fisher** proved that under the null hypothesis, $H_0: \rho = 0$, i.e., population correlation coefficient is zero.

The statistic\

$$t = \frac{r\sqrt{(n-2)}}{\sqrt{(1-r^2)}} \sim t_{(n-2)}$$ \hspace{1cm} [1.4.17]

If the value of 't' comes out to be significant, we reject $H_0$ at the level of significance adopted and conclude that $\rho \neq 0$, i.e., 'r' is significant of correlation in the population.

If 't' comes out to be non-significant then $H_0$ may be accepted and we conclude that variables may be regarded as uncorrelated in the population [61].

**Regression Analysis** ([51] & [52]):

The Regression analysis is a statistical device with the help of which we are in a position to estimate (or predict) the unknown values of one variable from known values of another variable. The variable which is used to predict the variable of interest is called the independent variable and the variable we are trying to predict is called the dependent variable.

**Definition:**

"Regression analysis attempts to establish the 'nature of the relationship' between variables, i.e., to study the functional relationship between the variables and thereby provide a mechanism for prediction or forecasting".

- Ya-Lun Chou
**Lines of Regression:**

The line of regression is the line which gives the best estimate to the value of one variable for any specific value of the other variable. Thus the line of regression is the line of 'best fit' and is obtained by the *principle of least squares.*

Let us suppose that in the bivariate distribution \((x_i, y_i); i = 1, 2, 3, \ldots n\). \(Y\) is dependent variable and \(X\) is independent variable.

Let,

the line of regression of \(Y\) on \(X\) be, \(Y = a + bX\), where \(a\) and \(b\) are constants.

The line of regression of \(X\) on \(Y\) be, \(X = A + BY\), where \(A\) and \(B\) are constants.

**Regression Equations:**

Regression equations, also known as estimating equations, are algebraic expressions of the regression lines. Since there are two regression equations viz.,

1) The regression equation of \(X\) on \(Y\)

\[ (X - \bar{x}) = r \frac{\sigma_x}{\sigma_y} (Y - \bar{y}) \quad \text{--------------------------} \quad [1.4.18] \]

2) The regression equation of \(Y\) on \(X\)

\[ (Y - \bar{y}) = r \frac{\sigma_y}{\sigma_x} (X - \bar{x}) \quad \text{--------------------------} \quad [1.4.19] \]

The regression equation of \(X\) on \(Y\) is used to describe the variations in the values of \(X\) for given changes in \(Y\) and the regression equation of \(Y\) on \(X\) is used to describe the variation in the values of \(Y\) for given changes in \(X\).
Properties of Regression Analysis:

1. Correlation coefficient is the geometric mean between the regression coefficients.
2. If one of the regression coefficients is greater than unity, the other must be less than unity.
3. Arithmetic mean of the regression coefficients is greater than the correlation coefficient (r), provided r>0.
4. Regression coefficients are independent of the change of origin but not on Scale.

Limitations of Regression Analysis:

In making estimate from a regression equation, it is important to remember that the assumption is being made that relationship has not changed since the regression equation had been computed. Another point worth remembering is that the relationship shown by the scatter diagram may not be the same if the equation is extended beyond the values used for computing the equation.

1.4. c) Markov Chain models [[3], [4], [24] & [54]]:

In fact, Stochastic Process is the mathematical description of a random phenomenon as it changes in time.

Definition

A Stochastic Process is a family of random variables \( \{ X_t \} \), where ‘t’ takes values in the index set ‘T’ (some times called a parameter set or time set). The values of \( X_t \) are called the state space and will be denoted by ‘S’.
If $T$ is countable then the Stochastic Process is called a *stochastic sequence* (or discrete parameter stochastic process). If $S$ is countable then the Stochastic Process is called a *discrete state process*.

If $S$ is a subset of the real line, the stochastic process is called a *real value process*.

If $T$ takes uncountable number of values continuously like $(0, \infty)$ or $(-\infty, \infty)$ the stochastic process is called a *continuous time process*.

### Different types of Stochastic Processes [54]:

The following are the most important types of stochastic processes we come across:

1. **Independent Stochastic Sequence (Discrete Time Process).**
   
   Here $T = [1, 2, 3, \ldots]$ and $\{X_t, t \in T\}$ are independent random variables.

2. **Renewal Process (Discrete Time Process).**
   
   Here $T = [0, 1, 2, 3, \ldots], S = [0, \infty]$. If $X_n$ are i.i.d. non-negative random variables and $S_n = X_1 + X_2 + \ldots + X_n$ then $\{S_n\}$ forms a discrete time renewal process.

3. **Independent Increment Process (Continuous Time Process)**
   
   Here $T = [t_0, \infty], \text{where } t_0 \text{ be any real number (positive or negative).}$
   
   For every $t_0 < t_1 < \ldots < t_n, t_i \in T, i = 1, 2, 3, \ldots, n \quad \text{----------[1.4.20]}$

   if $X_{t_0}, X_{t_1} - X_{t_0}, X_{t_2} - X_{t_1}, \ldots, X_{t_n} - X_{t_{n-1}}$ are independent for all possible choices of (1.4.20), then the stochastic process $\{X_t, t \in T\}$ is called *independent increment stochastic process*. 

38
4. Markov Process:

**Definition:**

A Stochastic system is called a Markov Process if the occurrence of a future state depends on the immediately preceding state and only on it. In other words future depends only on present and present depends only on the immediate past. This property is popularly known as 'Memory Less' property.

Here \( S = \text{a countable set, } T = \{0,1,2,3,\ldots\} \).

The Stochastic Process \( \{X_n, n=0,1,2,3,\ldots\} \) is called a Markov Chain, if, for \( j,k,j,\ldots,j_{n-1} \in N \) (or any subset of \( I \)),

\[
P\{X_n = k / X_{n-1} = j, X_{n-2} = j, \ldots, X_0 = j_{n-1}\} =
\]

\[
P\{X_n = k / X_{n-1} = j\} = p_{jk} \text{ (say)} \quad \text{[1.4.21]}
\]

whenever the first member is defined.

The outcomes are called the states of the Markov Chain, if \( X_n \) has the outcome 'j' (i.e. \( X_n = j \)), the process is said to be at state 'j' at \( n^{th} \) trial. To a pair of states \((j,k)\) at the two successive trials (say, \( n^{th} \) and \((n+1)^{th}\) trials) there is an associated conditional probability \( P_{jk} \). It is the probability of transition from the state 'j' at \( n^{th} \) trial to the state 'k' \((n+1)^{th}\) trial. The transition probabilities \( p_{jk} \) are basic to the study of the structure of the Markov Chain.

The transition probability may or may not be independent of 'n'. If the transition probability \( P_{jk} \) is independent of 'n', the Markov Chain is said to be homogeneous. If it is dependent on 'n', the chain is said to be non-
homogeneous. The transition probability $p_{jk}$ refers to the states $(j,k)$ at two successive trials (say, $n^{th}$ and $(n+1)^{th}$ trial); the transition is one-step and $p_{jk}$ is called **one-step transition probability**. In the more general case, we are concerned with the pair of states $(j,k)$ at two non-successive trials, say, state 'j' at the $n^{th}$ trial and state 'k' at the $(n+m)^{th}$ trial. The corresponding transition probability is then called **m-step transition probability** and is denoted by $p_{jk}^{(m)}$, 

i.e., $p_{jk}^{(m)} = \Pr\{X_{n+m} = k | X_n = j\}$. 

---------------- [1.4.22]

**5. Martingale or Fair Game Process** ([24], [25]):

If $E\left[\frac{X_{t_{n}}} {X_{t_{n}}} | X_{t_{n}} = a_{n}, X_{t_{n-1}} = a_{n-1}, \ldots, X_{t_{0}} = a_{0}\right] = a_{n}$

i.e. $E\left[\frac{X_{t_{n}}} {X_{t_{n}}, X_{t_{n-1}}, \ldots, X_{t_{0}}} \right] = X_{t_{n}}$ a.s. for all choices of the partition (1.4.20), then $\{X_{t}, t \in T\}$ is called a **Martingale Process**.

**6. Stationary Process:**

If the joint distribution of $(X_{t_{1}+h}, \ldots, X_{t_{n}+h})$ are the same for all $h > 0$ and $t_{1} < t_{2} < \ldots < t_{n}, t_{i} \in T, t_{i} + h \in T$ then $\{X_{t}, t \in T\}$ is called a **stationary process (strictly stationary process)**.

**Point process:**

When a countable set or sets of points randomly distributed on the real line or any arbitrary sets we call the family of random variables governed by the distribution of those random points as **point processes**.
**Transition Matrix**

The transition probabilities $p_{jk}$ satisfy $p_{jk} \geq 0$, $\sum_k p_{jk} = 1 \forall j$. 

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots \\ p_{21} & p_{22} & \cdots \\ \vdots & \vdots & \ddots \end{pmatrix} \quad \quad [1.4.23]$$

This is called the transition probability matrix (t.p.m.) of the Markov Chain. $P$ is a *Stochastic matrix*, i.e., a square matrix with non-negative elements and unit row sums. Apart from unit row sums if it has unit column sums also then $P$ is called *Doubly Stochastic Matrix*.

**Order of a Markov Chain:**

**Definition:**

A Markov Chain $\{X_n\}$ is said to be of order ‘s’ ($s = 1,2,3,...$), if, for all

$$\Pr\{X_n = k \mid X_{n-1} = j, X_{n-2} = j_1, \ldots X_{n-k} = j_{k-1}, \ldots\} = \frac{1}{s}$$

whenever the l.h.s. is defined.

**Transition Probabilities:**

**Chapman-Kolmogorov Equations** [24]:

$p_{jk}$ gives the probability of unit step transition from the state ‘j’ at a trial to the state ‘k’ at the next following trial. The m-step transition probability is denoted by

$$p_{jk}^{(m)} = \Pr\{X_{n+m} = k \mid X_n = j\}$$
gives the probability that from the state ‘j’ at $n^{th}$ trial, the state ‘k’ is reached at $(m+n)^{th}$ trial in m-steps, i.e. the probability of transition from state ‘j’ to the state ‘k’ in exactly ‘m’ steps.
Let $P = (p_{jk})$ denote the transition matrix of the unit-step transitions and $P^{(m)} = (p^{(m)}_{jk})$ denote the transition matrix of the $m$-step transitions. For $m = 2$, we have the matrix $P^{(2)} = P.P = P^2$. Similarly, $P^{(m+1)} = P^{(m)}P = P.P^{(m)}$ and

$$P^{(m+n)} = P^n P^m = P^m P^n.$$  

\[ \text{[1.4.25]} \]

**Classification of States and Chains:**

The states $j$, $j = 0, 1, 2, 3, \ldots$ of a Markov Chain $\{X_n, n \geq 0\}$ can often be classified in a distinctive manner according to some fundamental properties of the system. By means of such classification it is possible to identify certain types of chains.

**Communication Relations**

If $p^{(n)}_{ij} > 0$ for some $n \geq 1$, then we say that state $j$ can be reached or state $j$ is accessible from state $i$; the relation is denoted by $i \rightarrow j$. Conversely, if for all $n$, $p^{(n)}_{ij} = 0$, then $j$ is not accessible from $i$; in notation $i \not\rightarrow j$.

If two states $i$ and $j$ are such that each is accessible from the other then we say that the two states communicate; it is denoted by $i \leftrightarrow j$; then there exist integer $m$ and $n$ such that $p^{(n)}_{ij} > 0$ and $p^{(m)}_{ji} > 0$.

The relation $\rightarrow$ is transitive, i.e. if $i \rightarrow j$ and $j \rightarrow k$ then $i \leftrightarrow k$.

From Chapman-Kolmogorov equation

$$P^{(m+n)} = \sum_r P^{(m)}_r P^{(n)}_r$$

we get

$$P^{(m+n)}_i \geq P^{(m)}_j P^{(n)}_k.$$
Class Property:

A class of states is a subset of the state space such that every state of the class communicates with every other and there is no other state outside the class which communicates with all other states in the class. A property defined for all states of a chain is a class property if its possession by one state in a class implies its possession by all states of the same class. One such property is the periodicity of a state.

Periodicity

State ‘i’ is return state if \( p_i^{(n)} > 0 \) for some \( n \geq 1 \). The period \( d_i \) of a return to the state ‘i’ is defined as \( d_i = \text{G.C.D.}\{m: p_i^{(m)} > 0\} \); state ‘i’ is said to be aperiodic if \( d_i = 1 \) and periodic if \( d_i > 1 \). Clearly state ‘i’ is aperiodic if \( p_{ii} \neq 0 \).

Classification of Chains[24]:

If \( C \) is a set of states such that no state outside \( C \) can be reached from any state in \( C \), then \( C \) is said to be closed. If \( C \) is closed and \( j \in C \) while \( k \notin C \), then \( p_j^{(n)} = 0 \) for all, \( n \). i.e. \( C \) is closed iff \( \sum_{j \in C} p_j = 1 \) for every \( i \in C \).

A closed set may contain one or more states. If a closed set contains only one state ‘j’ then the state \( j \) is said to be absorbing; \( j \) is absorbing if and only if \( p_{jj} = 1, p_{jk} = 0, k \neq j \).

A Markov chain is irreducible if there is only one communicating class, i.e. if all states communicate with each other or every state can be reached from other state.
Classification of states [54]:

Transient and Persistent (Recurrent) States

Suppose that a system starts with the state \( j \). Let \( f_{jk}^{(n)} \) be the probability that it reaches the state \( k \) for the first time at the \( n \)th step and let \( p_{jk}^{(n)} \) be the probability that it reaches state \( k \) (not necessarily for the first time) after \( n \) transitions. Let \( \tau_k \) be the first passage time to state \( k \), i.e. \( \tau_k = \min\{n \geq 1, X_n = k\} \) and \( \{f_{jk}^{(n)}\} \) be the distribution of \( \tau_k \) given that the chain starts at state \( j \).

First Entrance Theorem

Whatever be the states \( j \) and \( k \),

\[
p_{jk}^{(n)} = \sum_{r=0}^{n} f_{jk}^{(r)} p_{jk}^{(n-r)}, \quad n \geq 1
\]

with \( p_{jk}^{(0)} = 1, f_{jk}^{(0)} = 0, f_{jk}^{(1)} = p_{jk} \)

First Passage Time Distribution

Let \( F_{jk} \) denote the probability that starting with state \( j \) the system will ever reach state \( k \). Clearly

\[
F_{jk} = \sum_{n=1}^{\infty} f_{jk}^{(n)} \quad \text{----------------------------- [1.4.26]}
\]

we have \( \sup_{n \geq 1} p_{jk}^{(n)} \leq F_{jk} \leq \sum_{n=1}^{\infty} p_{jk}^{(n)} \forall n \geq 1. \)

The mean (first passage) time from state \( k \) is given by

\[
\mu_{jk} = \sum_{n=1}^{\infty} n f_{jk}^{(n)} \quad \text{----- [1.4.27]}
\]

In particular, when \( k = j \), \( \{f_{jj}^{(n)}, n = 1, 2, \ldots\} \) will represent the distribution of the recurrence times of \( j \); and \( F_{jj} = 1 \) will imply that the return to the state \( j \) is certain. In this case

\[
\mu_{jj} = \sum_{n=1}^{\infty} n f_{jj}^{(n)} \quad \text{----------------------------- [1.4.28]}
\]

is known as the mean recurrence time for the state \( j \).
Definitions:

A state $j$ is said to be **persistent (or recurrent)** if $F_j^{-1} = 1$ (i.e. return to state $j$ is certain) and **transient** if $F_j < 1$ (i.e. return to state $j$ is uncertain). A persistent state $j$ is said to be **null persistent** if $\mu_j = \infty$, i.e. if the mean recurrence time is infinite, and is said to be **non-null (or positive) persistent** if $\mu_j < \infty$.

A persistent non-null and aperiodic state of a Markov chain is said to be **ergodic**.

Ergodicity

The behaviour in which sample averages formed from a process converge to some underlying parameter of the parameter of the process is termed ergodic. To make inference about the underlying laws governing an ergodic process, one need not observe separate independent replications of entire processes or sample paths. Instead, one need only observe a single realization of the process, but over a sufficiently long span of time. Thus, it is an important practical problem to determine conditions that leads to a stationary process being ergodic. The theory of stationary processes has a primary goal, the classification of ergodic behaviour and the prediction problem for process falling in the wide range of extremities.
**Ergodic Theorem:**

For a finite irreducible, aperiodic chain with t.p.m. \( P = (p_{ij}) \), the limits

\[
v_i = \lim_{n \to \infty} p_i^{(n)}
\]

exist and are independent of the initial state \( j \). The limits \( v_i \) are such that \( v_i \geq 0, \sum v_i = 1 \), i.e. the limits \( v_i \) define a probability distribution.

---

**1.4. d) Simulation Techniques** ([23], [26] & [57]):

A simulation is the imitation of the operation of a real-world process or system over time. It is a numerical technique for conducting experiments that involve certain types of mathematical and logical relationships necessary to describe the behaviour and structure of a complex real world system over an extended period of time.

Simulation is the process of generating values using random numbers without really conducting experiments. Experiments may be undertaken before the real system is operational so as to aid in its design, or to see how the system might react to change in its operating or to evaluate the system, response to change in its structure.

**When to Use Simulation**

The reasons for selecting simulation technique other than the known mathematical techniques are

(i) It is impossible to develop a mathematical solution.

(ii) Actual observation of a system may be too expensive.
(iii) Simulation may be the only method available because it is difficult to observe the actual environment.

(iv) There may not be a possibility to wait for a long period to study the system extensively.

(v) Actual operation and observation of a real system may be too disruptive.

**Methodology of Simulation:**

The methodology developed for simulation process consists of the following steps:

(i) Identifying and clearly defining the problem.

(ii) List the statement of objectives of the problem.

(iii) Construct an appropriate mathematical model of the given problem.

(iv) Ensure that the model represents the real situation.

(v) Make experiment with the model constructed, i.e., obtain a consistent set of values (or states) for the variables—a sample of what could happen in reality.

(vi) Analyse the results of simulation activity, i.e., use the sample obtained in above to calculate the value of the decision criterion, by actually following the relationships among the variables for each of the alternative decisions.

(vii) Make changes in the model, or the sample, and repeat the process until a sufficient number of samples are available.

(viii) Tabulate the various values of the decision criterion and choose the best policy.
**Simulation Models:**

A simulation model may be a physical or mathematical model, conceptual or a combination of them. Since physical models are relatively expensive to build, mathematical models are often preferred.

Broadly, the simulation models can be classified into the following four categories:

(i) **Simulation of Deterministic Models.** In the case of these models, the input and output variables are not permitted to be random variables and models are described by exact functional relationships.

(ii) **Simulation of Probabilistic Models.** In such cases, method of random sampling is used. The technique used for solving these models is termed as 'Monte-Carlo Technique'.

(iii) **Simulation of Static Models.** These models do not take variable time into consideration.

(iv) **Simulation of Dynamic Models.** These models deal with time varying interaction.

**Monte-Carlo Simulation** [14]:

Monte-Carlo method of simulation was developed by the two mathematicians John Von Neumann and S.M. Ulam. This method is generally used to solve problems which cannot be adequately represented by the mathematical models, or where the problems are too expensive for experimental solutions. This model involves random sampling from a known probability distribution and yields a solution which will be very close to the number of simulated trials which should be infinitive. Monte-Carlo simulation is generally computer oriented and performed as per steps given below.
**Step-1:** Identify the objectives of the problem and clearly define the problem.

**Step-2:** Construct an appropriate model and clearly specify the variables, parameters and the manner in which the time will change etc. Define the relationship between the variables and parameters.

**Step-3:** Prepare the model for experimentation by specifying the number of runs of simulation to be made.

**Step-4:** Determine the probability distribution for each variable in step 2 and establish the cumulative distribution function.

**Step-5:** Set up the table and assign tag numbers, with the help of cumulative distribution function.

**Step-6:** Generate random numbers and choose the corresponding tag number.

Then select the variable value corresponding to the tag number.

**Step-7:** Generate random numbers at many numbers of trials (given in the problem) and compute the values for different trials. The optimal solution will be the average values of different trials.

**Note:** *When the numbers of simulated runs are more, the obtained solution from this will be closer to the optimal.*

**Generation of Random Numbers** [[19], [23]]:

For solving all discrete problems, it is essential to generate random numbers. Most computer languages have a subroutine or a function that will generate random numbers. In carrying out Monte-Carlo simulation, one needs to generate random numbers to obtain random observations from a probability distribution. Random number is a sequence of numbers whose probability of
occurrence is the same as that of any other number in the sequence. The sequence of numbers must have two important Statistical properties, uniformity and independence. Each random number is an independent sample drawn from a continuous uniform distribution between 0 and 1. The probability density function (p.d.f.) is given by

\[
f(x) = \begin{cases} 
1, & 0 < x < 1 \\
0, & \text{otherwise}
\end{cases}
\]

Random numbers can be generated by any of the three methods.

(i) Manually generated random numbers.

(ii) Random numbers are selected from the random number table.

(iii) Computer-recursive algorithm.

Properties of Random numbers

1. Random numbers should be uniformly distributed.
2. It must be statistically independent.
3. They must have sufficiently long period.
4. The numbers must be generated at high speed.
5. They must occupy less memory.
6. They should not have a cyclic trend.

Tests for Random Numbers:

In literature, there are many tests available to test the desirable properties of Random Numbers uniformly and independently. Some of the important tests are given below. They are:
1. **Frequency test:** Uses the Kolmogorov-Smirnov or the chi-square test to compare the distribution of the set of numbers generated to a uniform distribution.

2. **Run test:** Tests the runs above or below the median by comparing the actual values to expected values.

   Let \( x_1, x_2, \ldots, x_n \) be the set of observations arranged in the order in which they occur. Then, for each of the observations, we see if it is above or below the value of the median of the observations and write 'A' if the observation is above and 'B' if it is below the median value. Thus we get a sequence of A's and B's of the type, (say), \( \text{ABBAAABABBBAAA....} \).

   **Definition of Run:**

   A run is defined as a sequence of letters of one kind surrounded by a sequence of letters of the other kind.

   Null hypothesis \( (H_0) \) that the set of numbers is random, the number of runs \('U'\) is a random variable with

   \[
   E(U) = \frac{n+2}{2} \quad \text{-----------------} \quad [1.4.30]
   \]

   and \[ Var(U) = \frac{n(n-2)}{2(n-1)} \quad \text{-----------------} \quad (1.4.31) \]

   For large \( n \) (say, \( > 25 \)), \( U \) may be regarded as asymptotically normal.

   Under \( (H_0) \) the test statistic is given by, \[
   Z = \frac{U - E(U)}{\sqrt{Var(U)}} \sim N(0,1) \quad \text{--[1.4.32]}
   \]

   If calculated \(|Z| \leq Z_\alpha\), we accept \( H_0 \) at \( \alpha \)% level of significance otherwise we reject \( H_0 \) and conclusion can be drawn accordingly.
Chi-Square ($\chi^2$) Test of Goodness of Fit:

A very powerful test for testing the significance of the discrepancy between theory and experiment was given by Prof. Karl Pearson in 1900 and is known as “Chi-Square test of goodness of fit”. It enables us to find if the deviation of the experiment from theory is just by chance or is it really due to the inadequacy of the theory to fit the observed data.

If $O_i, i = 1, 2, \ldots, n$ is a set of observed (experimental) frequencies and $E_i, i = 1, 2, \ldots, n$ is the corresponding set of expected (theoretical or hypothetical) frequencies, then Karl Pearson chi-square, given by:

$$\chi^2 = \sum_{i=1}^{n} \left[ \frac{(O_i - E_i)^2}{E_i} \right] \sim \chi^2(n - 2) \quad \text{[1.4.33]}$$

Where $\sum_{i=1}^{n} O_i = \sum_{i=1}^{n} E_i$

Conditions for the Validity of $\chi^2$ - test:

i. The sample observations should be independent.

ii. Constraints on the cell frequencies, if any, should be linear,

\[ \text{e.g., } \sum_{i=1}^{n} O_i = \sum_{i=1}^{n} E_i \]

iii. $N$, the total frequency should be reasonably large, say, greater than 50.

iv. No theoretical cell frequency should be less than 5. If any theoretical cell frequency is less than 5, then for the application of $\chi^2$ - test, it is pooled with the preceding or succeeding frequency so that pooled frequency is more than 5 and finally adjust for the degrees of freedom lost in pooling.
Advantages and Limitations of Simulation:

Advantages:

Simulation is sometimes described as indirect experimentation. In Operations Research we represent the system under study by constructing a model. Simulation is the process of experimenting on the model rather than on the operation which the model represents. Following are some of the advantages of Simulation:

i. The study of very complicated system or sub-system can be done with the help of simulation. Simulation has been described as 'what to do when all else fail'.

ii. By using simulation, we can investigate the consequences for a system of possible changes in parameters in terms of the model.

iii. The knowledge of a system obtained in designing and conducting the simulation is very valuable.

iv. It is a teaching aid, e.g., in business games, case studies, etc.

v. It enables us to assess the possible risks involved in a new policy before actually implementing it.

vi. The simulation of complicated systems helps us to locate which variables have the important influences on system performance.

vii. This can be used to experiment unfamiliar systems to prepare routine and extreme eventualities.

viii. Simulation methods are easier to apply than pure analytical methods.
Limitations of Simulation:

Use of simulation in place of other techniques, like everything else, involves a trade-off, and we should be aware of the disadvantages involved in the simulation approach. These include:

i. Simulation is not precise. It is not an optimization process and does not yield an answer but merely provides a set of the system's responses to different operating conditions. In many cases, this lack of precision is difficult to measure.

ii. A good simulation model may be very expensive. Often it takes years to develop a usable corporate planning model.

iii. Not all situations can be evaluated using simulation. Only situations involving uncertainty are considered, because without a random component, all simulated experiments would produce the same answer.

iv. Simulation generates a way of evaluating solutions but it does not generate the solution techniques. Managers must still generate all of the solution approaches they want to test.

v. Even when you spend the resources to build a simulation model that makes sense in some real-world context, it is often difficult for the people who built it to understand that they are still not looking at reality, but at best, an abstraction of the real world.

vi. Simulation is a time-consuming exercise.
Simulation is indeed a versatile tool. It provides only Statistical estimates rather than exact results and it only compares the alternatives rather than generating an optimal one. It is a slow and costly way to study a problem. Despite limitations it is an invaluable tool [19].

1.5. Chapter Summaries:

The Thesis consists of five chapters, the details of which are given below:

Chapter – 1: This chapter is basically an introductory one, where introduction is given to Cloud Seeding Experiments. The scope and limitations of main variable - rainfall (X), the concomitant variable - dam/reservoir level (Y), data on both the variables, Statistical tools and techniques used to analyze the data, are dealt with. Objectives of the thesis are also envisaged in a separate section. Finally, the chapter concludes with summaries of all the chapters of the Thesis.

Chapter – 2: This chapter concentrates on preliminary and basic analysis using ANOVA–One–Way classification. It is used to test the behaviour of Inflows into the dams in different rainy months. Based on the total monthly inflows into the dam, probability distribution is computed and finally the most probable rainy months are identified for Anantapur district. Conclusions are drawn based on the results obtained.
Chapter – 3: In this chapter, relation between the main variable - rainfall and the concomitant variable - dam/reservoir levels are calculated using Karl-Pearson correlation coefficients. Calculated correlation coefficients are tested for its significance, using Student’s-t test with appropriate degrees of freedom. Regression models are formed whenever the correlation is found significant and the rainfall is predicted using the proposed models. Using these predictions O.O.T. for conducting Cloud Seeding Experiments is determined.

Chapter – 4: In this chapter, a three state Markov Chain Model is proposed for the concomitant variable and various transition probabilities are calculated using the data collected and prepared one step transition probability matrices (t.p.m’s). Using these t.p.m’s, steady-state solutions are obtained for determining the behaviour of the concomitant variable. Using these steady-state probabilities, most probable rainy months are determined in a long run. After calculating Initial Probability Distributions (based on 2004 inflows), O.O.T. time for conducting Cloud Seeding Experiments in Anantapur district, is predicted.
Chapter – 5: This is the last chapter of this thesis where predictions are made based on Simulation Models. Using simulation technique, inflows into the dams are simulated using the steady-state probabilities obtained in Chapter–4. These simulated results are compared with actual results (based on the collected data, the closeness is tested through Chi-square test). Using these simulated models most probable rainy months are predicted and O.O.T. for conducting Cloud Seeding Experiments determined. The Chapter is finally concluded with further scope of the work.

The Thesis is finally appended with a list of References and Appendix-I, II & III, in the form a CD containing the original data collected on the main Variable (X) and Concomitant variable (Y).