VISVESVARAYA TECHNOLOGICAL UNIVERSITY

BELGAUM – 590 014

Shri. B. V. V. Sangha's

BASAVESHWAR ENGINEERING COLLEGE, BAGALKOT – 587 102

2009-2010

DEPARTMENT OF CIVIL ENGINEERING

A Project Report On

"GIS BASED RAINFALL AND RECHARGE STUDIES AND TIME SERIES ANALYSIS OF RAINFALL"

APPROVED BY KSCST, BANGALORE

PROJECT GUIDE
Dr. Veena Soraganvi

HEAD OF DEPARTMENT
Dr. C. B. Shivayogimath

PROJECT ASSOCIATES

Name of the Students	University No.
Ms. Deepa V. S.	*2BA06CV015*
Mr. Bhanuprakash B. Kulkarni	*2BA06CV013*
Mr. Niroop S. Chikkalli	*2BA06CV047*
Mr. Suraj G. Gadad	*2BA06CV084*

1. Introduction

Most of the rainfall in the region occurs during monsoon period. The rainfall received during the four months, i.e., June, July, August, September (JJAS) is considered as summer monsoon and is very crucial for the farming community. Researchers have used various approaches to study and predict the seasonal and intra-seasonal rainfall. To forecast the rainfall at different spatial and temporal scales, the models used in the past can be broadly classified as empirical and dynamical. This study is concerned with empirical models only. Empirical modeling strategies include identification of reliable precursors as well as optimal utilization of the information contained in the data on precursors.

The reasonable success achieved by the empirical approach has motivated persistent exploration of regional/global teleconnections of the Indian summer monsoon season since Walker's time, which resulted in a large number of predictors as well as a variety of statistical techniques. Because any modeling effort will have to be based on an understanding of the variability of the past data, ANNs have some special characteristics in this regard to be used. In contrast to conventional modeling approaches, ANNs do not require an in-depth knowledge of driving processes, nor do they require the form of the model to be specified a priori. This is true when modeling various climate variables for forecasting of hydrological variables like rainfall and stream flows. In view of this, ANNs are found to be a suitable approach for the prediction of Indian monsoon rainfall using large scale climate variables as input to the network. The next section provides a brief overview of climate variables and their influence on Indian monsoon rainfall.

1.1 Influence of Climate Variables on Indian Monsoon

At present, the assessment of the nature and causes of seasonal climate variability is still uncertain. There are still uncertainties associated with local and global climatic variables. For any rainfall prediction model, these are sources of variance in predictability. Recently, researchers have studied the influence and the possible relationships between various global climate variables and Indian monsoon rainfall. Additionally, they brought out several regional parameters based on sea-level pressure, temperature, and wind fields over India and sea surface temperature (SST) data from the adjoining Indian seas. Although their performance in seasonal forecasting has been encouraging, there is still a large variance in the monsoon rainfall unaccounted by the predictors identified so far. Several observational and modeling studies have indicated that the slowly varying surface boundary conditions, particularly in the winter and pre monsoon seasons, constitute a major forcing factor on the inter-annual variability of the monsoon rainfall. Parameters representing these conditions, global as well as regional, provide a handle for seasonal prediction. On inter annual timescales, the Indian monsoon rainfall has a strong and positive correlation with the pre-monsoon spring tropospheric temperature anomaly .Factors that influence the Indian summer monsoon include the sea surface temperature in the Pacific and Indian oceans global warming and human activities, among others.

1.2 Time Series Analysis

Time series analysis has become a major tool in forecasting and different applications in hydrology and environmental management fields. Among the most effective approaches for analyzing time series data is the model introduced by Box and Jenkins, ARIMA(Autoregressive Integrated Moving Average).In this study we used Box-Jenkins methodology to build ARIMA model for monthly rainfall data as well as yearly rainfall data. This research exclusively deals with time series forecasting model, in particular, the Auto Regressive Integrated Moving Average (ARIMA). These models series, for example rainfall, if these other time series data are correlated with a variable of interest and if there appears to be some cause for this correlation Box-Jenkins (ARIMA) modeling has been successfully applied in various water and environmental management applications. The followings are examples where time series analysis and forecasting are effective:

- **Water resources:** Time-series analysis has become a major tool in hydrology. It is used for building mathematical models to generate synthetic hydrologic records, to forecast hydrologic events, to detect trends and shifts in hydrologic record sand to fill in missing data and extend records.
- **Staff scheduling:** A manager of an environment department would need forecast of an hourly volume and type of waste generated to be processed in order to schedule staff and equipment efficiently.

- **Process control:** Forecasting can also be an important part of a process control system through monitoring key processes. It may be possible to determine the optimal time and extent of control action. For example, a chemical processing unit may become less efficient as hours of continuous operation increase.

1.2.1 ARIMA Models

The main stages in setting up a forecasting ARIMA model includes model identification, model parameters estimation and diagnostic checking for the identified model appropriateness for modeling and forecasting. Model Identification is the first step of this process. The data was examined to check for the most appropriate class of ARIMA processes through selecting the order of the consecutive and seasonal differencing required to make series stationary, as well as specifying the order of the regular and seasonal auto regressive and moving average polynomials necessary to adequately represent the time series model. The Autocorrelation Function (ACF) and the Partial Autocorrelation Function (PACF) are the most important elements of time series analysis and forecasting. The ACF measures the amount of linear dependence between observations in a time series that are separated by a lag k. The PACF plot helps to determine how many auto regressive terms are necessary to reveal one or more of the following characteristics: time lags where high correlations appear, seasonality of the series, trend either in the mean level or in the variance of the series. The general model introduced by Box and Jenkins includes autoregressive and moving average parameters as well as differencing in the formulation of the model. The three types of parameters in the model are: the autoregressive parameters (p), the number of differencing passes (d) and moving average parameters (q). Box-Jenkins model are summarized as ARIMA (p,d, q). For example, a model described as ARIMA (1,1,1) means that this contains 1 autoregressive (p)parameter and 1 moving average (q) parameter for the time series data after it was differenced once to attain stationary. In addition to the non-seasonal ARIMA (p, d, q) model, introduced above, we could identify seasonal ARIMA (P, D, Q) parameters for our data. These parameters are: Seasonal autoregressive (P), seasonal Differencing (D) and seasonal moving average (Q). For example, ARIMA (1,1,1) (1,1,1)12 describes a model that includes 1 autoregressive parameter, 1 moving average parameter, 1 seasonal autoregressive parameter and 1 seasonal moving average parameter. These parameters were computed after the series was differenced once at lag 1 and differenced once at lag 12.

1.3 Geographic Information Systems (GIS)

Geographic Information Systems (GIS) is an evolving, catch all phrase that initially referred to management of information with a geographic component primarily stored in vector form with associated attributes. This definition quickly became too restrictive with advances in software and ideas about information management. An advanced GIS system should be able to handle any spatial data, not just data tied to the ground by geographic reference points. The capacity to handle non-geographic spatial data was formerly the domain of systems referred to as AM/FM (Automated Mapping and Facilities Management). Other non-geographic applications, such as interactive medical encyclopedias that retrieve information based on the human form, should also be manageable by a robust system. Integration of imagery with vector data is now a necessity for a full-featured GIS system. Imagery was once thought to be the exclusive domain of image processing systems, but is now often required as a backdrop for vector, or other data, types. No up-to-date GIS system is complete without surface modeling and 3D (technically 2 1/2 D) visualization with "fly-by" capability. In addition to drawing a path for the simulation, you should be able to orbit with the view directed at a specified point or have the view pan around a stationary viewer. Vector overlay on this 3D surface should also be an integral part of the package.

A GIS system should be production oriented, which may or may not mean product oriented. Production work in GIS involves making maps (a product), but it also involves interactive analysis (a result which may have no tangible product). This booklet starts by looking at these two aspects of GIS systems and then describes the facets of GIS systems needed to reach the integrated goals.