# Analysis of rainfall data at Kurunagala A Stochastic approach

T. K. Hewapathirana
Department of Statistics & Computer Science, University of Kelaniya.

## ABSTRACT

Predication of rainfall in a particular region is very important for planning agricultural and other water related activities in that area. Therefore the present study was carried out to determine a suitable stochastic model for the prediction of the future rainfall pattern of the Kurunagala region of the Northwestern province of Sri Lanka. Monthly rainfall figures for Kurunagala from 1995-2002 were used for this analysis. Visual examination of the time plot showed that the variance of the series was not constant over time. As the variance was proportional to the mean of the data, the square root transformation was used in order to obtain a constant variance.

The correlogram of the transformed series gave evidence of a nonstationary mean together with a sinusoidal pattern. However, first order nonseasonal differencing was not sufficient to obtain a stationary mean. As strong seasonality can produce large autocorrelations at fractional seasonal lags first order seasonal differencing was used to check whether the nonstationarity was due to this fact.

The series obtained was found to be stationary with a correlogram exhibiting a peak at lag 12 followed by a cut off after the first seasonal lag. The partial correlogram exhibited a significant coefficient at the first seasonal lag together with a dying down pattern at the seasonal lags. This fact was clear in the rainfall figures from 1990-2002 as well. Hence an ARIMA$(0,0,0)(0,1,1)_{12}$ model was fitted to the data. After estimating the model parameters by non-linear least squares the model was found to be

$$(1 - B^{12})\sqrt{x_t} = (1 - 0.9\,B^{12})a_t$$

where

$x_t$ - observation at time t
$a_t$ - random stock at time t
B - backward shift operator

The normal probability plot and the histogram of the residuals were found to be satisfactory. The autocorrelation coefficients and the partial autocorrelation coefficients too were within the 95% confidence limits. The Ljung-Box Chi-Square statistics were found to be nonsignificant and the residual sum of squares of the fitted model were found to be low, giving evidence of a satisfactory model. Although exact predictions were not possible due to the high variability in the data. It was found that all historic data values fall within the 95% confidence limits of the predicted values.