Volume-5
Issue - I
June-2020

Rajarata University Journal

ISSN: 2362-0080

www.ruj.ac.lk/journals/

Category: Research Article

Fitting a Poisson-Gamma Model for Rainfall Occurrence and Rainfall Amount in Colombo District, Sri Lanka

*Selvaratnam, Ahthiga & Wijekoon, Pushpakanthie

Department of Statistics and Computer Science, Faculty of Science, University of Peradeniya, Peradeniya (20400), Sri Lanka

ARTICLE DETAILS

Article History

Published Online: 30 June 2020

Keywords

Gamma distribution, Generalized linear models, Poisson distribution, Tweedie distribution, Zero-inflated

***Corresponding Author**

ahthigas@gmail.com

ABSTRACT

Rainfall modelling plays an important role in many sectors in Sri Lanka such as agriculture and hydrology, for predicting and forecasting rainfall. Most of the studies on modelling rainfall are based on statistical and mathematical techniques. The variables that are considered for rainfall modelling in most of the studies are rainfall occurrence and rainfall amount. Since the rainfall occurrence is a discrete variable, the Poisson distribution and the method of Markov chain are the widely used methods for modelling. The gamma distribution, exponential distribution and mixed exponential distribution are the common distributions applied to model the rainfall amount. In this study, an attempt has been made to fit a single distribution to model both rainfall occurrence and rainfall amount simultaneously, by using generalized linear models with family of Tweedie distribution. Daily rainfall amounts in Colombo district from 1951 to 2000 are used for model fitting and the relevant data from 2001 to 2007 are used for model validation, respectively. To study the cyclical patterns and seasonal variations of daily rainfall amount, a model with six-monthly frequency sine-cosine terms is fitted. The model fits well to the data based on the mean daily rainfall amounts for 50 years. This model produces continuous values indicating the rainfall events for which exact zero values indicate the non occurrence of rainfall. Simulated rainfall data for the period from 2001 to 2007 resembles the actual data in terms of summary statistics. This model can be further improved by adding more predictors such as climate indices.

1. Introduction

Since the economy of Sri Lanka significantly depends on its agriculture, it is very important to study the patterns of rainfall. Most of the farmers in Sri Lanka use techniques inherited from their ancestors to identify the seasonal variations and predict weather patterns [1]. However, with the global climate change their predictions have become less reliable. Drought is one of the main issues in Northern and Eastern province of Sri Lanka [2]. The drastic drop in the water level in major reservoirs causes widespread crop failure and food and drinking water shortages. The main reason for this situation is failure of rain on time. On the other hand, heavy rains and flood become major threats to lives. Landslide threats are major problems in the live hoods of people of Central province. The recent disasters that occurred in Sri Lanka have raised the importance of improving the studies of rainfall modelling to provide much useful information to natural disaster management, water resources management, agricultural and hydrological sectors.

In most of the studies, rainfall occurrence and rainfall amount are the variables that are considered for rainfall modelling. Since the rainfall occurrence is a discrete variable, the Poisson distribution and the method of Markov chain are the widely used methods for modelling. The gamma distribution, exponential distribution and mixed exponential distribution are the common distributions applied to model the rainfall amount. Hence, rainfall modelling requires two distributions to be fitted. Even though these models are easy to understand and interpret, there are many parameters to be estimated and several assumptions to be made in carrying out the whole process. Moreover, rainfall amount data may consists of excess number of exact zeros, which make it zero-inflated. Usually hourly and daily rainfall data are zero-inflated compared to monthly, seasonally and annual rainfall data, because of the short duration. Thus, modelling rainfall amount with exact zeroes is very important.

A number of researchers have used Markov chain models and time series analysis to study rainfall in Sri Lanka. In 2016, Ariyarathna and Alibuhtto [3] used a univariate time series seasonal autoregressive integrated moving average model for modelling monthly rainfall in Katunayake region. They fitted SARIMA $(2,0,2)(2,0,1)_{12}$ model for monthly rainfall data for the period January 2001 - January 2016.

In 2005, Jayawardene et al. [4] focused on analyzing the trends and seasonal patterns of rainfall on a set of 15 selected weather stations. They presented their work by computing Mann-Kendall rank statistic, Spearman rank statistic and regression parameters on annual rainfall data from 1869 to 1980. They analysed the data for a long term period(1869 onwards i.e 98 to 130 years) and for a short term period(1949 onwards i.e 36 to 50 years).

In 2012, in the study of Sonnadara [5], the first order Markov chain probability models were used to study the wet and dry spells of observed at the Colombo meteorological station (1941-2000) based on daily precipitation.

Dunn [6] is the first researcher who succeeded in producing a model for rainfall occurrence and rainfall amount, that requires only one distribution (2003). He used Tweedie distribution to model both discrete (non-rainfall occurrence/ zero rainfall amount) and continuous (rainfall occurrence/ positive rainfall amount) components of rainfall. In 2004, he modelled monthly precipitation at Charleville, Australia using Tweedie distribution for each month. He also used Tweedie distribution to model daily and monthly precipitation in Melbourne, Australia. Later, several researches carried out the same process of fitting Tweedie distribution for rainfall modelling on daily and monthly time scale.

In 2010, Hasan and Dunn [7] used a simple Poisson-gamma model for modelling rainfall occurrence and rainfall amount simultaneously. They used monthly rainfall data from 220 Australian rainfall stations with 6 stations as case studies. They simulated monthly rainfall amounts for 96 years using their final model, to be used in agricultural production systems.

In 2018, Dzupire et al. [8] developed a single model for both rainfall occurrence and rainfall amount. In this study they have used daily rainfall data of Balaka district from 1995-2015 period and simulated rainfall amounts for 2015 and 2016.

The objective of this article is to use a single distribution known as Tweedie distribution to represent both occurrence and amount of rainfall in Colombo district. Specifically, to identify the appropriate generalized linear model in family of Tweedie distribution to model daily rainfall occurrence and amount simultaneously.

This article is structured as follows. Section 2 presents methods for formulating the model and estimation of the parameters of the model. Preliminary analysis of the data, the results obtained by fitting a model to the estimation data, diagnostic checks and evaluation of the final model are given in Section 3. Finally, the study is concluded in Section 4.

2. Material and Methods

a. Poisson-gamma Distribution

In this article, we are interested in using the Poisson distribution and Gamma distribution for modelling rainfall occurrence and rainfall amount respectively. A Poisson-gamma distribution or a Compound Poisson distribution is the sum of independent and identically gamma distributed N random variables, where N follows a Poisson distribution. Dzpuire et al. [8] have given a framework for deriving the parameters of the Poisson-gamma model using the parameters of Tweedie distribution. A brief explanation of mathematical formulation of the model is given as follows.

b. Formulation of the Model

Let the number of rainfall occurrences on a particular day say N, follows a Poisson distribution with mean λ. The probability mass function of N is given as:

$$P(N=n) = e^{-\lambda}\frac{\lambda^n}{n!}, \text{for all } n \in \mathbb{N} \qquad (1)$$

Assuming the amount of rainfall recorded for the $i^{th}(1 \le i \le N)$ rainfall event in that day, say y_i, follows a Gamma distribution with shape parameter α and rate parameter γ such that $y_i \sim Gamma(\alpha, \gamma)$. Then the probability density function of y_i can be given as:

$$f(y) = \begin{cases} \frac{\alpha^{\gamma} y^{\gamma-1} e^{-\alpha y}}{\Gamma(\gamma)} & , y > 0 \\ 0 & , y \leq 0 \end{cases} \quad (2)$$

Let R be the total sum of the rainfall amount of each rainfall event in that day and it is given by eq. (3). Since $y_1, y_2, \ldots, y_N$ are independent and identically distributed gamma random variables, R has a gamma distribution with the parameters $N\alpha$ and γ. Then R is assumed to be gamma distributed and independent of the number of rainfall occurrences, N, on that day.

$$R = y_1 + y_2 + \cdots + y_N \; for \; N > 0 \quad (3)$$

If $N = 0$, then $R = 0$. Therefore, the probability of recording no rainfall on a particular day is

$$P(R = 0) = e^{-\lambda} \frac{(\lambda)^0}{0!} = \exp(-\lambda) \quad (4)$$

Hence, the random variable R follows a mixed distribution with mass at 0.

Here R is the sum of random variables $y_1, y_2, \ldots, y_N$ from a gamma distribution where N is an independent Poisson random variable. Therefore, the distribution of R given N is a compound Poisson distribution or Poisson-gamma distribution. Thus, the cumulant generating function of R can be given as follows:

$$\ln M_R(s) = \lambda(M_Y(s) - 1) = \lambda[(1 - \alpha s)^{-\gamma} - 1] \quad (5)$$

Dzupire et al. [8] showed that λ and α cannot be expressed in closed form using the likelihood function derived from the probability density function of R. Therefore, they have expressed the probability density function in terms of exponential dispersion models.

An exponential dispersion model (EDM) has the probability density function of the form

$$f(y; \theta, \phi) = a(y, \phi) \exp\left\{\frac{1}{\phi}[y\theta - k(\theta)]\right\} \quad (6)$$

for suitable functions $a()$ and $k()$. Here y is the random variable of interest, $\phi > 0$ is the dispersion parameter and θ is the canonical parameter. The function $k(\theta)$ is the cumulant of the exponential dispersion model.

Let a random variable Y follows an EDM with mean μ and dispersion parameter ϕ such that $Y \sim ED(\mu, \phi)$. The mean of the distribution is expressed as $\mu = E(Y) = k'(\theta) = dk(\theta)/d\theta$ and the variance is given as $Var(Y) = \phi k''(\theta) = \phi(d^2k(\theta)/d\theta^2)$. The relationship between μ and θ given by $\mu = dk(\theta)/d\theta$ is invertible, so that θ can be expressed as a function of μ. Therefore, the variance can be expressed as $Var(Y) = \phi V(\mu)$, where $V(\mu)$ is called a variance function. To obtain the variance function $V(\mu)$, Tweedie family distributions can be used. Tweedie family distributions are special case of EDMs which is defined as below.

The family of exponential dispersion models with variance function of the form $V(\mu) = \mu^p$ for $p \notin (0,1)$, are called Tweedie family distributions. The mean of the distribution is μ and the variance is $\phi\mu^p$. Here p is called the index parameter.

The following are some special cases of the Tweedie distributions:

- When $p = 0$, it gives the Normal distribution.
- When $p = 1$ and $\phi = 1$, it gives the Poisson distribution.
- When $p = 2$, it gives the Gamma distribution.
- When $p = 3$, it gives the Inverse Gaussian distribution.
- When $1 < p < 2$, the Tweedie distribution corresponds to the Poisson-gamma distribution.

By using the variance function of Tweedie family distribution, the variance of a random variable Y which follows Tweedie EDMs with mean μ and dispersion parameter ϕ, such that $Y \sim Tw_p(\mu, \phi)$, can be written as:

$$Var(Y) = \phi\mu^p \; ; \; p \notin (0,1) \quad (7)$$

The Tweedie distribution densities cannot be expressed in closed form for values of p except for $p = 0, 1, 2$ and 3. Therefore, Tweedie distribution for $1 < p < 2$ does not have a closed form of density function. Instead, it can be identified using its cumulant generating function.

For $1 < p < 2$, the cumulant generating function of a Tweedie distribution is given as follows.

$$\log M_Y(t) = \frac{1}{\phi}\frac{\mu^{2-p}}{2-p}\left[\begin{matrix}(1 + t\phi(1-p)\mu^{p-1})^{(2-p)/(1-p)} \\ -1\end{matrix}\right] \quad (8)$$

By comparing the cumulant generating functions in eq. (5) and eq. (8), parameters of the Poisson-gamma distribution $(\lambda, \alpha, \gamma)$ can be expressed in terms of parameters of the Tweedie distribution (μ, ϕ, p) as:

$$\lambda = \frac{\mu^{2-p}}{\phi(2-p)}$$

$$\alpha = \phi(p-1)\mu^{p-1} \quad (9)$$

$$\gamma = \frac{2-p}{p-1}$$

Based on Tweedie distribution, the probability of recording no rainfall on a day is $P(R=0) = \exp\left[-\frac{\mu^{2-p}}{\phi(2-p)}\right]$ and the probability of having a rainfall event is

$$P(R>0) = W(\lambda,\alpha,R,\gamma)\exp\left[\frac{R}{(1-p)\mu^{p-1}} - \frac{\mu^{2-p}}{2-p}\right], \quad (10)$$

where $W(\lambda,\alpha,R,\gamma)$ is an example of Wright's generalized Bessel function expressed as $W(\lambda,\alpha,R,\gamma) = \sum_{j=1}^{\infty}\frac{\lambda^{j}(\alpha R)^{j\gamma}e^{-\lambda}}{j!\Gamma(j\gamma)}$.

c. Estimation of Parameters

The Tweedie family of distributions is a family of EDMs with three parameters namely mean μ, dispersion parameter ϕ and index parameter p. The response variables of generalized linear models(GLMs) follow EDMs. Therefore, fitting a Tweedie distribution follows the framework of fitting a GLM. Dunn and Smyth [9] have defined a generalized linear model as given below.

A generalized linear model can be defined as follows:

i. Independent response variables $Y_1, Y_2, \cdots, Y_n$ are generated such that $Y_i \sim ED(\mu_i, \phi_i/w_i)$ where w_i are known prior weights.
ii. The relationship between the means μ_i and linear predictors is provided by a monotonic link function g, so that $g(\mu_i) = x_i^T\beta$ where x_i is a vector of covariates and β is a vector of unknown regression parameters.

The log link function is selected as the most suitable link function for Poisson-gamma distribution. Hence, the models of the form $log(\mu_i) = X_i\vec{\beta}$ can be fitted where $\mu_i \sim Tw_p(\mu,\phi)$. In order to fit the Tweedie family, the estimates of β, p and ϕ should be found.

The estimation of parameters can be done in R software using the *tweedie* package created by Dunn [9] [10] [11]. For any given value of p, a robust iterative procedure called iteratively reweighted least-squares is used to estimate β. The maximum likelihood estimate(MLE) of p can be found using profile likelihood technique which requires an iterative procedure using computer computations. The function *tweedie.profile* is used to estimate the MLE of p.

First, a sequence of possible values of p is generated from 1.2 to 1.8 with an increment of 0.1. Then for any value of p in the generated sequence, the maximum likelihood estimate of β and ϕ are found and the log-likelihood is computed. This procedure is repeated for all possible values in the sequence. Thus, a cubic-spline interpolation is plotted through these computed points, which is known as a profile likelihood plot. The value of p which maximizes the log-likelihood is selected as the MLE $\hat{p}$. The estimated values of regression coefficients and ϕ can be obtained by fitting a generalized linear model with family of Tweedie distribution for $\hat{p}$ with a log link function.

3. Results and Discussion

3.1. Data Collection and Preliminary Analysis

The rainfall data were collected from Department of Meteorology, Colombo. Records include the daily rainfall amount (in millimetre) in Colombo from 1st of January 1951 to 31st of December 2007. If the rainfall amount is zero in a particular day, then that day is a dry day which has not experienced any rainfall events.

The rainfall data set has 10787 days (51.81%) of no rainfall days out of 20819 days (57 years) recorded. Summary statistics shows that the minimum and maximum values are 0mm and 493.7mm respectively. The mean rainfall amount for the whole data set is 6.58mm. There were 30 missing values out of 20819 daily rainfall amounts and they were estimated using the mean weightings method. Out of 57 years of data, daily rainfall amounts for the first 50 years (1951 - 2000) were used for the parameter estimation of the model while the rest (2001 - 2007) were used for validation purposes.

3.2. Model Fitting

To understand the nature of data, the relationship between the mean and the variance of the data should be identified. Therefore, the variance and the mean of the rainfall amount for each day of the year were computed. Then the log of the variance was plotted against the log of the mean which is shown in figure 1. The plot shows that there is an approximate linear relationship between log(mean) and log(variance). It also shows that there are some extreme points. This relationship can be expressed as:

$$\log(Variance) = \alpha + \beta\log(mean) \quad (11)$$

$$Variance = \exp\{\alpha + \beta\log(mean)\} \quad (12)$$

$$Variance = constant \times mean^{\beta} \quad (13)$$

That is, the variance can be given as some power $\beta \in \mathbb{R}$ of the mean. This agrees with the variance function of the Tweedie family distributions $V(\mu) = \mu^p\,; p \notin (0,1)$.

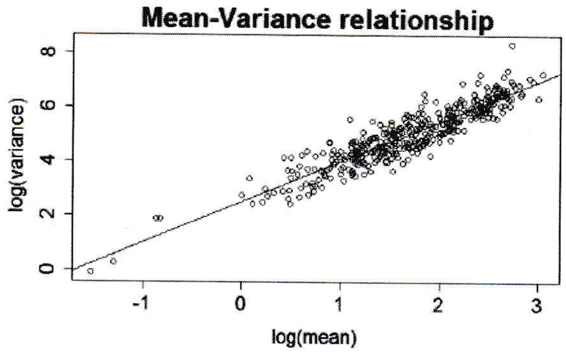

Figure 1. The mean–variance relationship of daily rainfall amounts

A model with six monthly frequency sine and cosine terms as predictors was fitted to model the cyclic nature and seasonality of rainfall throughout a year. We assumed that the month of February in all the years in the data set has 28 days. Hence, the log link function can be given as:

$$log(\mu_i) = \beta_0 + \beta_1 sin\left(\frac{4\pi i}{365}\right) + \beta_2 cos\left(\frac{4\pi i}{365}\right), \quad (14)$$

where $\mu_i \sim Tw_p(\mu, \phi)$ and β_0, β_1 and β_2 are regression coefficients. Here $i = 1,2, \ldots, 365$ corresponds to days of a year that is Julian day.

The plot of the profile log-likelihood function was used to find the maximum likelihood estimate of index parameter $\hat{p}$. The value of p that maximizes the profile likelihood function (L) is 1.628571 and the maximum value of L is - 43195.39.

Using the estimated value of p, a generalized linear model was fitted to the estimation data set. Table 1 gives the results obtained by fitting the generalized linear model. The results show that the maximum likelihood estimate of the dispersion parameter $\hat{\phi}$ is 12.07935 and the coefficients of the six monthly sine and cosine terms are highly significant, and thus they are identified as important characteristics of rainfall.

The predicted mean rainfall amount of each day of the year can be given by the following equation.

$$\hat{\mu}_i = exp\left(1.76460 - 0.66103\, sin\left(\frac{4\pi i}{365}\right) - 0.30810\, cos\left(\frac{4\pi i}{365}\right)\right) \quad (15)$$

A comparison between actual mean and predicted mean of daily rainfall amount is given by the plot in figure 2. Using this plot, we can say that the predicted mean resembles the actual mean of daily rainfall amount.

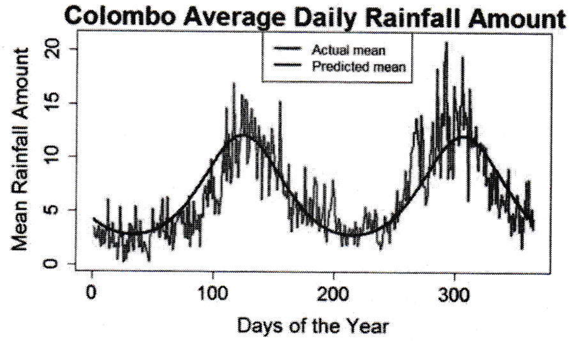

Figure 2. Actual versus predicted mean rainfall amount

Hence, the fitted model was found to be performing well in predicting mean daily rainfall amount and the coefficients of the covariates were highly significant. This indicates that the six monthly frequency sine and cosine terms are important characteristics for modelling rainfall in Colombo district.

The parameters of Poisson distribution and Gamma distribution were estimated using the estimated index parameter $\hat{p}$ and the estimated dispersion parameter $\hat{\phi}$ of Tweedie distribution. They are given as:

$$\hat{\lambda}_i = \frac{1}{4.49112}\left[exp\left(1.76460 - 0.66103\, sin\left(\frac{4\pi i}{365}\right) - 0.30810\, cos\left(\frac{4\pi i}{365}\right)\right)\right]^{0.37143} \quad (16)$$

$$\hat{\alpha}_i = 7.56012\left[exp\left(1.76460 - 0.66103\, sin\left(\frac{4\pi i}{365}\right) - 0.30810\, cos\left(\frac{4\pi i}{365}\right)\right)\right]^{0.62857} \quad (17)$$

$$\hat{\gamma} = 0.59091 \quad (18)$$

Using eq. (16), the probability of no rainfall occurrence for Julian days of the year can be calculated as:

Table 1. Estimated parameter values of the fitted model

Parameter	Estimate	Std. Error	*t* value	*Pr(>|t|)*
$\hat{\beta}_0$	1.76460	0.01871	94.33	< 2e-16 ***
$\hat{\beta}_1$	-0.66103	0.02630	-25.14	< 2e-16 ***
$\hat{\beta}_2$	-0.30810	0.02613	-11.79	< 2e-16 ***
$\hat{\phi}$	12.07935			

Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

$$\Pr(R_i = 0) = \exp\left(-\frac{1}{4.49112}\left[exp\left(1.76460 - 0.66103\, sin\left(\frac{4\pi i}{365}\right) - 0.30810\, cos\left(\frac{4\pi i}{365}\right)\right)\right]^{0.37143}\right) \quad (19)$$

By using eq. (19), the probability of receiving no rainfall for each day of the year was calculated. The calculated probability was plotted against days of the year. The plot is given in figure 3.

Figure 3 shows that the probability of no rainfall occurrence follows a seasonal pattern. It indicates that the probability is high if the previous days have not experienced any rainfall events and vice versa. This agrees with the principle of the Markov process in rainfall modelling.

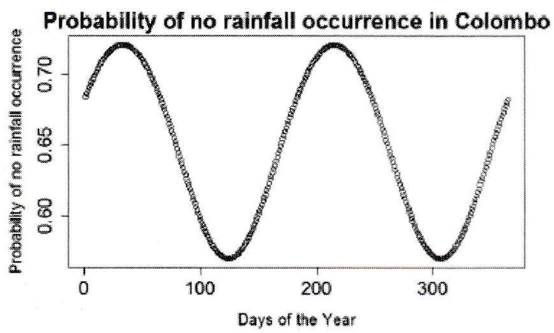

Figure 3: Probability of no rainfall occurrence for days of the year

3.3. Diagnostic Check

The main diagnostic available for GLM is Q-Q plot of the quantile residuals. If the quantile residuals are normally distributed, then the correct distribution was chosen for the mean of the daily rainfall amount and the fitted model would perform well. The normal Q-Q plot for the fitted model is given in figure 4(1).

The normal Q-Q plot shows that the residuals in the centre part lie on the straight line. The residuals in the upper tail show a small deviation. However, the residuals in the lower tail show a large deviation from the straight line compared to the residuals in the upper tail. Therefore, the normality of the quantile residuals cannot be checked using this plot and it requires a normality test to be performed.

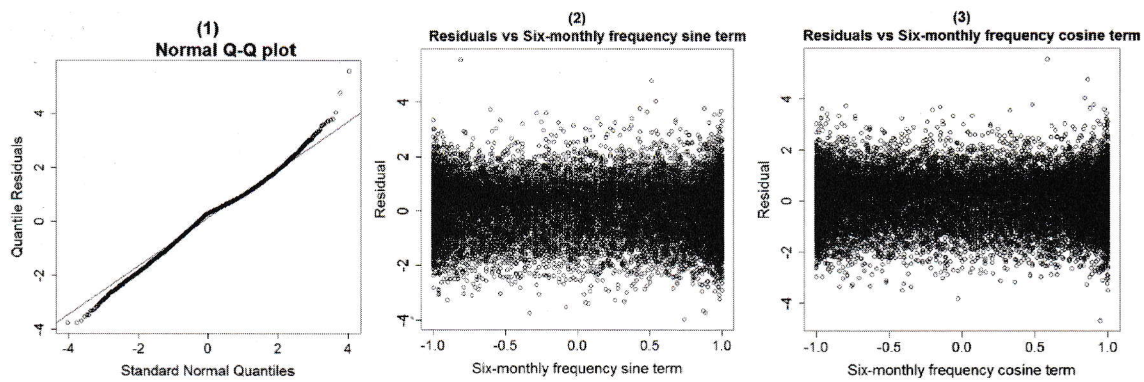

Figure 4: Normal Q-Q plot of the quantile residuals and residuals versus predictors for the fitted model

Since this is a non-normal regression and the data set is zero-inflated, none of the normality tests are appropriate. Therefore, the plot of residuals versus covariates were used to examine the distribution of the response variable and to check whether the fitted model is appropriate or not Figure 4(2) and figure 4(3) show the plots of residuals versus covariates for covariates six-monthly frequency sine term and six-monthly frequency cosine term respectively. Since the plots do not show any pattern, we can suggest that the fitted model is adequate and there is no need for adding or using different sets of predictors in the model.

3.4. Model Evaluation

In order to determine the predicting ability of the final model, it was simulated for a period of 7 years. Daily rainfall amounts from 2001 to 2007 were used for the model validation. A random sequence of rainfall amounts with a length of 2555 (number of days in 7 years) was drawn using Tweedie distribution with parameters based on the estimated value of mean rainfall amount for each days of the year $\mu_i ; i = 1,2, \ldots, 365$, the estimated index parameter $\hat{p}$ and the estimated dispersion parameter $\hat{\phi}$. By changing the number seeds, various random sequences were obtained.

However, they do not show a significance difference in between them.

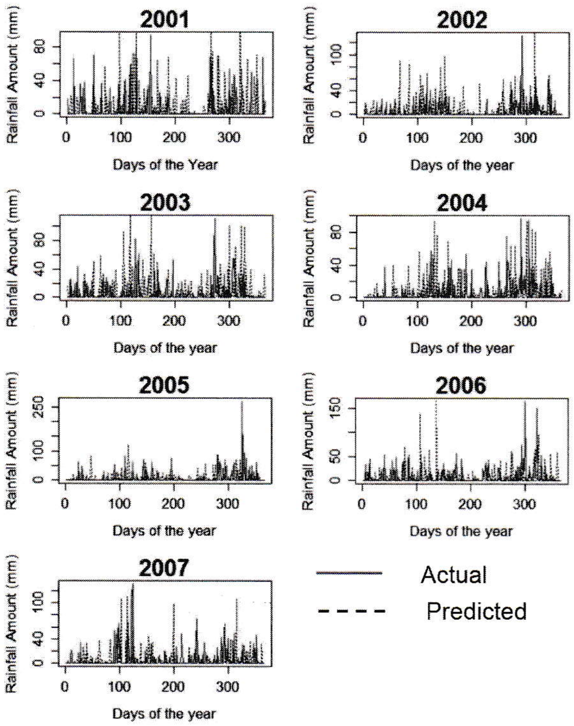

Figure 5: Actual versus predicted daily rainfall amount for the period 2001-2007

The daily rainfall amounts simulated from the final model were compared with the actual data using the summary statistics. For the actual data, the mean, median, IQR and maximum values (in mm) are 6.123, 0.0, 4.4 and 270.1 respectively. The mean, median, IQR and maximum values (in mm) of the simulated data are 6.631, 0.0, 4.0, and 251.0 respectively. The comparison shows that the simulated data satisfactorily predicts the overall mean, median, IQR and maximum values of daily rainfall amounts. There were 1692 days and 1342 days with no rainfall occurrences (zero rainfall amount) in the simulated data and actual data respectively. The fitted model over estimated the number of days with no rainfall occurrences. The time series plot of the actual versus the predicted daily rainfall amount for the period of 2001 to 2007 is given in figure 5.

From figure 5 we can see that the simulated rainfall amount pattern for 2005, 2006 and 2007 shows the resemblance to the pattern of the actual rainfall amount of the respective years. However, for the years 2001 - 2004, the actual and simulated data patterns do not show a significant resemblance to each other.

4. Conclusion

The main purpose of this study is to understand the characteristics of Tweedie distribution to be utilized in rainfall modelling. Unlike the previous studies in Sri Lanka related to rainfall modelling using two different distributions for rainfall occurrence and rainfall amount, this study acts as an initiative approach to fit a single distribution called Tweedie family distribution which can represent both components of rainfall in Sri Lanka. The ability to model exact zeros is an added advantage of this model.

A model with six-monthly frequency sine and cosine terms as predictors was developed using daily rainfall amounts in Colombo district. Even though this study has demonstrated a way to fit Tweedie family distribution to single site, modelling rainfall in multisite using this distribution is also possible (Swan [12]; Hasan and Dunn [7]). Therefore, this model can be further developed to model rainfall in other districts as well.

The model can be fitted to different time scales such as monthly which is very easy to fit and interpret. Further improvements can be done by adding more appropriate predictors such as Southern Oscillation Index, Southern Oscillation Phase and Indian Ocean indices to forecast rainfall in agricultural and hydrogeology sectors.

References

1. Hirimathugodage D. Sri Lankan Farmers' Traditional Knowledge and Climate Change Predictions. [Internet] http://climatenet.blogspot.com/2018/07/sri-lankan-farmers-traditional.html (Accessed on: 12/03/2019)
2. ACAPS Briefing Note: Sri Lanka - Drought [Internet]. [Available from: https://reliefweb.int/sites/reliefweb.int/files/resources/acaps_start_briefing_note_sri_lanka_drought_170223.pdf (Accessed on: 15/03/2019)
3. Ariyarathna PAHR, Alibuhtto MC. Modeling monthly rainfall in Katunayake region using seasonal ARIMA model. Enriching the Novel Scientific Research for the Development of the Nation 2016; 82-96.
4. Jayawardene HKWI, Sonnadara DUJ, Jayewardene DR. Trends of rainfall in Sri Lanka over the last century. Sri Lankan Journal of Physics 2005; 6:7-17.
5. Sonnadara DUJ. Modeling daily rainfall using Markov chains. *Annual* Research Symposium, University of Colombo 2012; 327-329.

6. Dunn PK. Occurrence and quantity of precipitation can be modelled simultaneously. International Journal of Climatology 2004; 24:1231–1239.

7. Hasan MM, Dunn PK. A simple Poisson-gamma model for modeling rainfall occurrence and amount simultaneously. Agricultural and Forest Meteorology 2010; 150:1319–1330.

8. Dzupire NC, Ngare P, Odongo L. A Poisson-gamma model for zero inflated rainfall data. Journal of Probability and Statistics 2018; 1012647:12 pages.

9. Dunn PK, Smyth GK. Series evaluation of Tweedie exponential dispersion model densities. Statistics and Computing 2005; 15(4):267-280.

10. Dunn PK, Smyth GK. Evaluation of Tweedie exponential dispersion models using Fourier inversion. Statistics and Computing 2008; 18(1):73-86.

11. Dunn , Smyth GK. Tweedie: Evaluation of Tweedie exponential family models. R package version 2.3 2017.

12. Swan T. Generalized estimating equations when the response variable has a Tweedie distribution: An application for multi-site rainfall modeling, The University of Southern Queensland, Toowomba, QLD, 2006.

