Noh N.M., Atan I., Jaafar J., Ashaari Y., Samsudin M.B., Khatib M.R., Baki A. (2021) A simple approach to estimate rainfall at different return period, pp. 73-80. In Gastescu, P., Bretcan, P. (edit, 2021), *Water resources and wetlands*, 5th International Hybrid Conference Water resources and wetlands, 8-12 September 2021, Tulcea (Romania), p.235
Available online at http://www.limnology.ro/wrw2020/proceedings.html Open access under CC BY-NC-ND license
5th International Hybrid Conference Water resources and wetlands, 8-12 September 2021, Tulcea (Romania)

A SIMPLE APPROACH TO ESTIMATE RAINFALL AT DIFFERENT RETURN PERIOD

Nurbaiah Mohammad NOH[1], Ismail ATAN[1], Jurina JAAFAR[1], Yasmin ASHAARI[1], Mohd Bakri SAMSUDIN[2], Mohd Riza KHATIB[2] and Aminuddin BAKI[2,3*]
[1]Faculty of Civil Engineering, UiTM, 40450, Shah Alam, Selangor, Malaysia
[2]Envirab Services, PO Box 7866, GPO Shah Alam 40730, Selangor, Malaysia
[3]Faculty of Engineering, City University Malaysia, Menara CityU, Petaling Jaya, Selangor, Malaysia.
*Email: aminbaki@msn.com

Abstract. The estimation of rainfall of different return period is very important in designing hydraulic structures of the area. Most of the available methods require the application of trial and error procedure to determine the suitability of distribution function for any particular section. The method suggested in the present paper can be easily applied without testing the suitability of different distribution functions. The Johnson transformation was known to be able to reduce the coefficient of skewness to near zero for rainfall data. The transformed rainfall data are fitted to the normal distribution and used in the design of hydraulic structures. The study indicates that the transformed normal distribution can be used as the determine approach towards estimating rainfall at different return period. Rainfall estimation is an important tool in water resources planning and management.
Keywords: coefficient of skewness, Johnson transformation, normal distribution, rainfall data

1.INTRODUCTION

Most analysis is for rainfall of various durations at a single station. For each duration provides data which are analyzed for frequency and combined into a family of curves. Recording rain gages provide the only satisfactory data for frequency analysis for rainfalls of short duration.

Though number of models such as, normal, two-parameter log normal, three-parameter log normal, Person III, Gumbel and log Gumbel etc need to be tested for arriving at the suitable model for estimation of rainfall amount of different return period at any station, they required computation of various parameters which required much time and becomes difficult to compute.

Numbers of transformation functions have already been tested for the a few places in Penang and Federal Territory Kuala Lumpur however, none of them was as convenient as the family of power transformation. With these viewpoints, an attempt was made to decide an alternative method by considering estimated rainfall by Power transformation as the observed rainfall at different return periods. Johnson transformation can be used to analyse daily rainfall data of different stations of Penang and Federal Territory Kuala Lumpur.

The estimation of rainfall of different return period is very important in designing hydraulic structures of the area.

2.LITERATURE REVIEW

Studies by Kuczera (1983) stated the power normal transformation distribution is similar to the log-Pearson distribution in the sense that the transformation parameter λ is akin to the log-Pearson distribution in skew parameter. Furthermore, because of the tractability of the power normal transformed distribution, the inferences of λ using site and regional data can be studied analytically.

Siriwardena *et al.* (2002) found Srikanthan and McMahon (1985a, b) measured that the rainfall values in the last class are modelled by a Power transformation and the values in the intermediate classes are modeled by a linear distribution. In a subsequent modification, shifted Gamma distribution is used to model the last class. Srikanthan and McMahon (2001) observed monthly probabilities of rainfall for each weather type were used to classify a day as wet or dry. The rainfall amounts were modelled using the product of an Exponential random variable and a Uniform random variable because an Exponential distribution alone underestimated the variance of the daily rainfall.

Studies by e.g., Chander *et al.* (1978); McAully (1979), Jain and Singh (1987); Lye (1992) and Kuczera (1983) stated that in hydrologic modelling, the power normal transformation has been found to perform well in a number of at site flood frequency studies as well as in regional flood frequency analysis. Xu *et al.* (2019) utilised transformation in their studies of inflows. Wang *et al.* (2020) used data transformation in their studies of river flow. Barber (2020) studied goodness of fit of hydrological data using data transformations.

All those studies used in depth statistical analysis, which can be referred to in Box and Cox (1964), Shapiro and Francia (1972), Filliben (1975), Hinkley (1977), Taylor (1983), Looney and Gulledge (1985), D'Agostino *et al.* (1999) and Mishra (1999). Strong hydrological knowledge is also essential Subramanya (1984).

3.MATERIALS AND METHODS

3.1 The study area

Daily rainfall data for 5 different satiation namely Rumah Kebajikan Pulau Pinang, Klinik Bukit Bendera, T/Air Besar Sg. Pinang, Ibu Bekalan KM 16 and Sek. Men. Taman Maluri of Penang and Wilayah Persekutuan for the period 22, 22, 26, 14 and 14 years respectively were collected from Department of Irrigation and Drainage Malaysia (DID).

3.2 Data analysis and transformation data

The moving totals were computed to identify 1-D (1 day) to consecutive 7-D (7 days) maximum rainfall for each year for all selected places. The data transformed to normality by using Johnson Transformation function, which obtain from the probability plot. The data become normal when the coefficient of skewness and kurtosis were near to zero. Different statistical parameter, such as, mean (M); standard deviation (S); the coefficient of variation (C_v); skewness coefficient (G_1); and kurtosis coefficient (G_2) for historical and transformed data were computed using the software (Johnson transformation) developed in minitab.

For example, the rainfall station at Rumah Kebajikan was selected based on the following reason:

- The null hyphothesis underlaying normal distribution was rejected for all different extremes (1-D to 7-D maximum).

Hence, the data was selected in order to decide the efficacy of the transformed normal distribution. The rainfall amount at different return period for 5-yr, 10-yr, 20-yr, 50-yr, 100-yr, 200-yr, 500-yr and 1000-yr was estimated and retransformed to its original

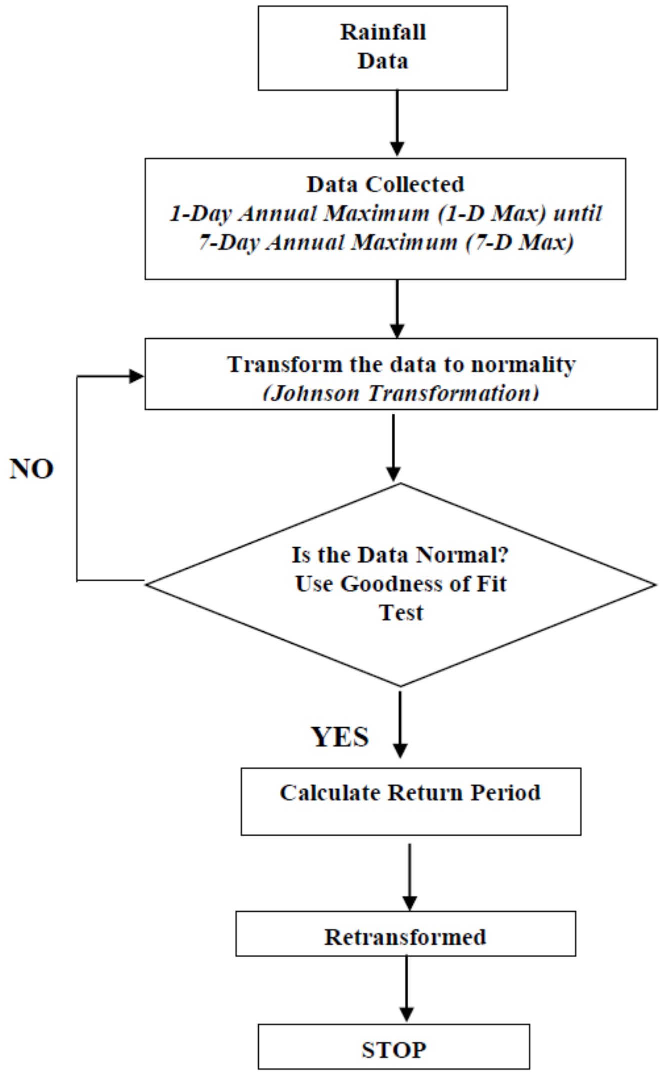

Figure 1 Schematic diagram of the research

4.RESULTS AND DISCUSSION

4.1 Transform the data to normality

The data become normal when the coefficient of skewness and kurtosis were near to zero. The P-Value (0.021) is less than 0.1, the null hypothesis is rejected. The results are statistically significant and the data are normally distributed. The transformed data, y in computed by the transformation function, $y = -6.4415 + 1.5503 * \text{Log}(x - 29.8694)$ (refer to figure 1.1). However, the result from the transformation for station Ibu bekalan KM 16 showed that all the data does not to transform to normality because of the null hypothesis underlying normal distribution was not rejected for all different extremes, so no transformation required.

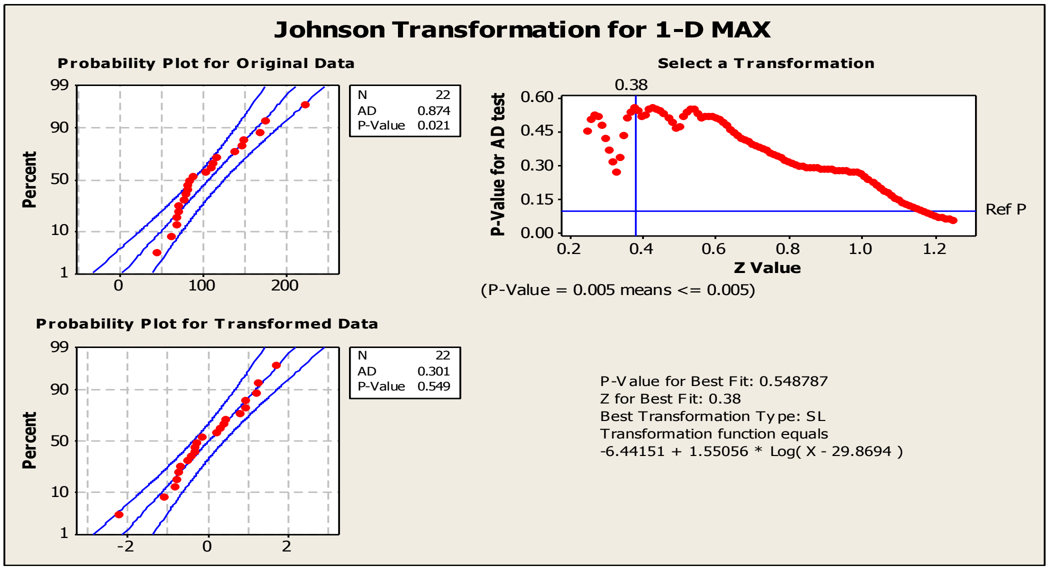

Figure 1. Probability Plot for Historical data and Transformed data at Rumah Kebajikan

4.2 Variation of Statistical Parameters in Johnson Transformation

Before and after transformation, some of the statistical parameters are computed e.g. mean (M); standard deviation (S); the coefficient of variation (C_v); skewness coefficient (G_1); and kurtosis coefficient (G_2). The normality can be achieved by using this transformation if the skewness and kurtosis is near to zero is had shown in Table 1.

Table1: Estimated rainfall amounts for different extremes and return periods (X_T) for Rumah Kebajikan by Johnson Transformation method for different statistical parameters of model

X_T,years	1-D	2-D	3-D	4-D	5-D	6-D	7-D
5	135.8	171.6	188.5	203.1	224.5	256.4	255.2
10	167.6	226.9	257.8	278.5	286.1	302.5	306.8
20	201.1	294.7	359.4	385.4	345.9	349.8	360.7
50	248.5	407.0	560.0	590.8	415.9	398.9	437.1
100	287.1	511.9	776.7	808.0	459.8	426.8	499.0
200	328.4	636.0	1067.8	1092.7	495.7	448.6	565.6
500	387.7	835.1	1593.6	1597.0	532.8	469.8	660.1
1000	435.9	1015.0	2126.6	2097.0	554.1	481.6	737.3
M	0.0	0.0	-0.1	-0.1	0.0	0.1	0.0
S	0.9	1.0	0.8	0.9	1.0	1.0	1.0
C_v	12038.7	3258.3	-1322.0	-1202.6	-6707.6	1951.2	2820.6
G_1	-0.2	-0.3	-0.2	-0.3	-0.3	-0.1	-0.2
G_2	0.0	-0.3	-0.9	-0.7	-0.3	0.0	-0.4

4.3 Estimation of Rainfall amounts from Johnson Transformation

An attempt was made to forecast the rainfall amounts of different return periods for Rumah kebajikan, as a critical station, based on the probability plot after Johnson transformation function for all extremes was linear in the range of 12% to 95% of probability levels.

The return period for 5-yr, 10-yr, 20-yr, 50-yr, 100-yr, 200-yr, 500-yr and 1000-yr are computed by using the frequency factor for use in normal distribution. The values of frequency factor use in normal distribution corresponding to return period, X_T, years shown in **table 2.**

Table 2: Frequency Factor use in Normal Distribution corresponding to Return Period, X_T, years

.**Years, X_T**	**Standard normal deviate, *t***
5	0.8416
10	1.2816
20	1.6449
50	2.0538
100	2.3264
200	2.5758
500	2.8782
1000	3.0902

The values of return period, X_T can be calculated by using mean and standard deviation from the transformed data and the equations as below:

Return Period,

$$X_T = \mu_y + t\sigma_y$$

Where,
μ_y = mean, t = standard normal deviate, σ_y = standard deviation

The actual return period was obtained by transforming the data back to its original scale and the resulted of the return period for all stations, X_T are shown in Table 3.

Table 3: estimated rainfall amounts for different extremes and return periods (X_T) for T/Air Besar Sg. Pinang, Klinik Bukit Bendera, Rumah Kebajikan, Ibu Bekalan KM 16 and Sek. Ren. Taman Maluri

T/AIR BESAR SG. PINANG AT P.PINANG							
years	**1-D**	**2-D**	**3-D**	**4-D**	**5-D**	**6-D**	**7-D**
5	179.8	242.0	287.6	307.5	322.8	340.3	355.8
10	211.1	268.9	319.5	340.8	358.0	376.9	391.4
20	242.3	291.2	345.9	368.4	387.1	407.1	420.8
50	284.4	316.2	375.5	399.4	419.8	441.1	453.9
100	317.4	332.9	395.3	420.0	441.6	463.7	475.9
200	351.4	348.1	413.3	438.9	461.6	484.4	496.1
500	398.7	366.6	435.3	461.9	485.8	509.6	520.5
1000	436.0	379.6	450.6	477.9	502.7	527.2	537.7
KLINIK BKT. BENDERA at P.PINANG							
5	172.3	226.5	266.8	301.5	308.9	333.7	387.2
10	207.7	310.2	335.2	377.5	406.8	432.4	431.4
20	246.2	435.5	419.2	467.1	544.0	564.3	467.9
50	303.0	687.1	558.8	610.7	802.4	803.9	509.0

100	351.1	964.1	689.6	740.5	1072.9	1045.2	536.4
200	403.7	1337.1	845.2	891.0	1421.5	1348.2	561.5
500	481.4	2019.5	1095.1	1126.1	2035.5	1865.6	591.9
1000	546.4	2720.9	1322.4	1335.1	2641.4	2361.0	613.2
RUMAH KEBAJIKAN at P.PINANG							
5	135.8	171.6	188.5	203.1	224.5	256.4	255.2
10	167.6	226.9	257.8	278.5	286.1	302.5	306.8
20	201.1	294.7	359.4	385.4	345.9	349.8	360.7
50	248.5	407.0	560.0	590.8	415.9	398.9	437.1
100	287.1	511.9	776.7	808.0	459.8	426.8	499.0
200	328.4	636.0	1067.8	1092.7	495.7	448.6	565.6
500	387.7	835.1	1593.6	1597.0	532.8	469.8	660.1
1000	435.9	1015.0	2126.6	2097.0	554.1	481.6	737.3

IBU BEKALAN KM.16 at GOMBAK, W.PERSEKUTUAN							
5	102.7	135.2	159.3	177.6	188.4	202.1	216.2
10	108.3	144.5	171.2	191.5	203.1	215.7	230.0
20	112.9	152.1	181.0	203.1	215.1	226.9	241.3
50	118.2	160.8	192.0	216.0	228.7	239.4	254.1
100	121.6	166.5	199.4	224.7	237.8	247.8	262.6
200	124.8	171.8	206.1	232.6	246.1	255.5	270.4
500	128.7	178.2	214.3	242.2	256.1	264.8	279.8
1000	131.4	182.6	220.0	249.0	263.2	271.3	286.5
SEK.REN.TAMAN MALURI at W.PERSEKUTUAN							
5	111.2	131.0	167.3	189.1	203.0	231.2	256.6
10	119.0	153.8	196.6	221.7	245.8	295.9	318.5
20	125.4	188.4	235.9	257.8	298.6	383.0	391.7
50	132.6	259.5	307.9	311.4	387.3	541.0	508.6
100	137.4	339.2	380.9	357.1	470.4	700.1	614.9
200	141.8	448.6	473.5	407.9	570.1	901.1	738.2
500	147.1	653.1	633.2	482.9	731.1	1244.1	931.1
1000	150.8	864.9	787.6	546.1	877.5	1573.8	1101.9

5.CONCLUSION

Based on the findings of the study, the following conclusions are drawn:

- The Johnson Transformation is capable of transforming daily rainfall data to normality.
- The behavior of rainfall extremes for different stations can be described by the Johnson transformation.

- This transformation reduces the range of the variable due to which probability function become linear and prediction that is more accurate can be made.
- Method of Johnson transformation is perhaps to solve the problems of choosing a distribution when the region has a short record.
- Avoid the problem of choosing the suitable distribution was by transforming the data to normality using Johnson Transformation.
- It is an easy approach in estimate rainfall of any return periods.
- The estimated rainfall amounts for different stations can be utilized for design of hydraulic structures for soil and water conservation and drainage of excessive runoff from the field crops.

ACKNOWLEDGEMENT

Appreciation to the Department of Irrigation and Drainage (JPS) Malaysia for making their data available for this study.

REFERENCES

Barber, C. A. (2020). Evaluating Statistical Goodness-of-Fit Estimators for Skewed Hydrologic Data (Doctoral dissertation, Tufts University) *unpublished*.

Box, G.E.P. and Cox D.R. (1964), An Analysis of Transformation, *Journal of the Royal Statistical Society, 26,* 211-252

Chander, S., Spolia, S.K. and Kumar, A. (1978), Flood frequency analysis by power transformation, *ASCE Journal of the Hydraulics Division,* Vol. 104 (HY11), 1495-1504

D'Agostino, R.B., Belanger, A. and D'Agostino, R.B., Jr. (1990), A suggestion for using powerful and informative tests of normality, *The American Statistician,* Vol. 44(4), 316-321

Filliben, J.J. (1975), The probability plot correlation coefficient test for normality, *Technometrics,* Vol. 17(1), 111-117

Hinkley, D. (1977), On quick choice of power transformation, *Applied Statistics,* 26, 67-69

Jain, D., & Singh, V. P. (1987). Comparison of some flood frequency distributions using empirical data. In *Hydrologic Frequency Modeling* (pp. 467-485). Springer, Dordrecht.

Kuczera, G. (1983). Improved parameter inference in catchment models: Combining different kinds of hydrologic data and testing their compatibility. *Water Resources Research*, *19*(5), 1163-1172.

Looney, S.W. and Gulledge, T.R. (1985), Use of the correlation coefficient with normal probability plots, *The American Statistician,* Vol. 39(1), 75-79

Lye, L. M. (1992), A technique for selecting the Box-Cox transformation in flood frequency analysis, *Faculty of Engineering and Applied Science*, 760-766

McAully R.S. (1979), Flood Frequency Analysis by Power Transformation, *Water Survey of Canada*, 369-381

Mishra. K K (1999), Statistical Analysis of Rainfall Data, *Journal of the Institution of Engineers*, Vol. 71

Shapiro, S.S. and Francia, R.S. (1972), An approximate analysis of variance test for normality, *Journal of the American Statistical Association,* Vol. 67, 215-216

Srikanthan, R., & McMahon, T. A. (1985a). Gould's probability matrix method: 1. The starting month problem. *Journal of Hydrology*, *77*(1-4), 125-133.

Srikanthan, R., & McMahon, T. A. (1985b). Gould's probability matrix method: 2. The annual autocorrelation problem. *Journal of Hydrology*, *77*(1-4), 135-139.

Srikanthan, R., & McMahon, T. A. (2001). Stochastic generation of annual, monthly and daily climate data: A review.

Siriwardena, L., McMahon, T. A., & Srikanthan, R. (2002). Evaluation of two daily rainfall data generation models. *Cooperative Research Centre for Catchment Hydrology*.

Subramanya K (1984), Hydrology, *Tata McGraw,* New Delhi, Edition 2

Taylor, M.J.G. (1983), Power Transformation to symmetry, Journal of the American Statistical Association, Vol. 81, 114-118

Xu, W., Fu, X., Li, X., & Wang, M. (2019). Data transformation models utilized in Bayesian probabilistic forecast considering inflow forecasts. *Hydrology Research*, *50*(5), 1267-1280.

Wang, Q. J., Bennett, J. C., Robertson, D. E., & Li, M. (2020). A data censoring approach for predictive error modelling of flow in ephemeral rivers. *Water Resources Research*, e2019WR026128.