Estimation of Probable Maximum Precipitation in the Context of Climate Change

KMSS Gunathilaka[1#], WCDK Fernando[1], and SS Wickramasuriya[1]

[1]Faculty of Engineering, *General Sir John Kotelawala Defence University, Ratmalana, Sri Lanka*

[#]<sumudumalie.sathsaranie@gmail.com>

Abstract- *Probable Maximum Precipitation (PMP) estimates are essential when designing hydraulic structures, especially since the risk of the failure of such structures is high. The impact on climate change with PMP has been crucial at present although the concept does not incorporate climate change. Although there are two widely used methods to estimate PMP, this research focused on the statistical method, covering 16 stations of the Kelani River catchment. The daily precipitation records for 57 years were collected, and the annual maximum daily rainfall series was prepared for all 16 stations. The study was conducted using five scenarios (S1- S5). The results from Hershfield PMP (S1) emphasize that the Hershfield enveloping curve has a very high value of frequency factor (K) in low annual average maximum daily precipitation. Thus, the need to modify the curve has arisen as a major objective of this research. As a result, Modified Hershfield PMP (S2) and Modified PMP in the context of Sri Lanka (S3) are considered. Outlier detection (S4) manifests that, there may be one or more or devoid of outliers deviating from the original concepts of Hershfield. Split sampling (S5) concludes that Standard Deviation (SD) is the most influential factor for PMP, which shows the effect of climate change. PMP maps are developed to observe the spatial-temporal variation of PMP, which is the first version in the context of Sri Lanka.*

Keywords— **PMP, Statistical method, Modified enveloping curve, Split sampling, Climate change.**

I. INTRODUCTION

The evolution of dams from then, till now has gathered several engineering technologies with retrospect from past experiences especially, dam failures. Dam safety is the most crucial section that draws attention to dam design (Silva and Premakumar, 2011). The high risk of dam failures is due to the inundation of downstream causing tremendous losses of lives and infrastructure. Thus, dam failures should be mitigated as much as possible to reduce the consequences. Probable Maximum Flood (PMF) is used to design the spillway structure of major dams where the risk of failure is very high. PMP is the input in estimating the PMF. Kelani River catchment has several dams that have been built and Moragahakanda, Laxapana, Upper Kothmale, and Victoria, are few dams in Kelani River catchment. So, the need to estimate Probable Maximum Precipitation (PMP) for the Kelani River catchment was raised.

PMP is defined as the "greatest depth of precipitation for a given duration meteorologically possible for a design watershed or a given storm area at a particular location at a particular time of year, with no allowance made for long-term climatic trends" (WMO, 2009).

Although the definition of the PMP does not account for long-term climatic trends, WMO has included the effects of climatic change in the manuals of 2009 considering its importance in the current context. It is argued that climate change does not affect directly to PMP unless there is an observed climatic increase or decrease (WMO, 2009). According to the Sri Lankan context, there are no clear patterns or trends identified in precipitation (Ministry of Mahaweli Development and Environment, 2016). However, some regions in South Africa and some urban regions in India show increasing rainfall, opposing the above statement (Carter, 1996). Therefore, analyzing the changes in climate with PMP is crucial.

Hydro-meteorological method and the statistical method are two widely used methods to estimate PMP. If sufficient precipitation data are available and other meteorological data such as wind records and dew point temperatures are lacking, it is convenient to use the statistical approach developed by Hershfield which focuses on the frequency analysis and influence of outliers (WMO, 2009; Thanh Thuy, Kawagoe and Sarukkalige, 2019). Considering the availability of data, a statistical approach has been used for PMP estimation in this study. Daily rainfall data from 1960-2016 (57 years) were collected from 16 rainfall stations covering the Kelani River catchment area.

II. METHODOLOGY

The study has been conducted using 5 scenarios.

A. Scenario 1 (S1) - Hershfield Statistical Method

The daily rainfall data were preliminary screened and prepared annual maximum daily rainfall series for all stations. The mean and the standard deviation of the series were calculated with and without the maximum observed rainfall event.

Hershfield statistical method (1961a, 1961b) which is explained in the WMO Manual emphasizes that there are adjustment factors for the standard deviation and the mean due to the maximum observed event and the length of record (WMO, 2009). Accordingly, adjustments were calculated and according to the Hershfield method, the maximum observed event was always an outlier.

Hershfield used the general frequency equation (see Equation 1) developed by Chow to estimate PMP (WMO, 2009).

$$X_m = \bar{X}_N + (K \times S_N) \qquad \textit{Equation 1}$$

Where X_m = maximum observed rainfall, $\bar{X}_N$ = mean of the annual maximum daily rainfall for N years for the particular station, K = common statistical variable and S_N = standard deviation. Also, the frequency factor K_m can be calculated using the equation (Sarkar and Maity, 2020c).

$$K_M = \frac{X_m - \bar{X}_{(N-1)}}{S_{(N-1)}} \qquad \textit{Equation 2}$$

Where, K_M = the frequency factor, X_m = Annual maximum daily rainfall of a particular station, $\bar{X}_{(N-1)}$ and $S_{(N-1)}$ are the respective mean and the standard deviations excluding the maximum observed rainfall event.

B. Scenario 2 & 3 (S2/ S3) -Historical Data Analysis and Modification of the Hershfield Enveloping Curve

There is no exact universal method to estimate the PMP values. The conventional method of estimating the K value is by using the Hershfield method. The frequency factor (K) plays a major role in the journey of estimating the PMP value. High values of K lead to an overestimation of PMP causing an uneconomical design while the low value gives an underestimation following the design structure at high risk (Wickramasuriya and Fernando, 2012; Sarkar and Maity, 2020a). Therefore, the K value should be calculated with care.

The K values in the enveloping curve developed by Hershfield depict a difference with real scenarios especially in low average annual maximum daily precipitation (Sarkar and Maity, 2020b, 2020c). Thus, the need to modify the curve is enhanced. In this study, one of the major objectives of the research was to obtain a modified curve considering historical extreme rainfall data from the Kelani River catchment and outside the catchment.

Historical extreme values of the years 1940 and 1947 were obtained for 10 stations in the Kelani River catchment area which are considered under scenario 2 (S2). According to the recorded maximum value, the Frequency factor K (Historical Kelani) was calculated as per and a graph was plotted with average annual maximum daily precipitation vs K to see the behaviour of the enveloping curve with the actual data set.

Similarly, Historical data sets outside the catchment were obtained for a particular period with extremely high rainfalls and the frequency factor K (Historical Sri Lanka) was gleaned under scenario 3 (S3). The historical K values of each aspect were plotted on a graph with annual average maximum daily precipitation vs K to identify the enveloping value for the Kelani catchment and the Sri Lankan context.

Sri Lanka is a small island covering a small area, so the rainfall pattern does not change much spatially(Bank, 2021). Therefore, the enveloping curve was modified according to the maximum K value obtained from the two options (S2 & S3), which is Km. K value was modified as Km for the low average annual maximum daily rainfall (mean values lower than $\bar{X}$m). But the values higher than the mean value of $\bar{X}$m are not modified in scenario 2.

In scenario 3, the whole enveloping curve was modified even after the ($\bar{X}$m) to get a conclusion for estimating a PMP value for an optimized design satisfying the main aim of this study.

C. Scenario 4 (S4) - Detection of Outliers

Outlier detection test was conducted using the Interquartile Range Test (IQR). This method was chosen due to its simplicity and does not need prior knowledge about the distribution of the data set. This is a non-parametric test done with the concepts of quartiles and interquartile ranges.

Initially, the first (Q_1) and third (Q_3) quartiles of the data series of the Annual maximum daily precipitation of each station are calculated and the interquartile range is reckoned.

$$IQR = (Q_3 - Q_1) \qquad \textit{Equation 3}$$

The threshold value to detect the outliers is stated in the following equation.

$$P_{thresh} = Q_3 + (3 \times IQR) \qquad \textit{Equation 4}$$

Where P_{thresh} = Threshold precipitation limit for the outlier detection, IQR = interquartile range given from IQR = (Q3 – Q1) Equation 3 and Q_3 = third quartile of the dataset of Annual maximum daily precipitation. These outliers obtained from P_{thresh} = Q3 + (3 × IQR) Equation 4 are termed as extreme outliers while the mild outliers are obtained from the equation $P_{thresh} = Q_3 + (1.5 \times IQR)$ (Walfish, 2006; Fernando and Wickramasuriya, 2021). However, this research only focuses on the extreme outliers. After knowing the P_{thresh}, it is needed to identify the number of outliers in each station. In the presence of more than one outlier, the adjustments for the mean and the standard deviation may vary from the Hershfield method. If the P_{thresh} is in between X n and X (n-1) for a data series of X_1, X_2, X_3.........$X_{(n-2)}$, $X_{(n-1)}$ and X_n the ratios for the adjustment factors would be $\frac{\bar{X}(N-1)}{\bar{X}(N)}$ and $\frac{S(N-1)}{S(N)}$. But for a case of two outliers, the P_{thresh} lying before $X_{(N-1)}$ the ratio must be $\frac{\bar{X}(N-2)}{\bar{X}(N)}$ and $\frac{S(N-2)}{S(N)}$ where, $\bar{X}_{(N-2)}$ and $S_{(N-2)}$ are the respective mean and the standard deviation of the data set after excluding both outliers. Also, these values follow the pattern of, $\frac{\bar{X}(N-1)}{\bar{X}(N)} > \frac{\bar{X}(N-2)}{\bar{X}(N)}$ and $\frac{S(N-1)}{S(N)} > \frac{S(N-2)}{S(N)}$ (Fernando and Wickramasuriya, 2021). For the stations with no outliers the original mean and standard deviations are used without any adjustments. The PMP was obtained for all the values after identifying the outliers for the modified K value.

D. *Scenario 5 - (S5) Split sampling*

The data set of annual maximum daily precipitation of 57 years was split into two samples the sample set from 1960-1987 and the other sample as 1988-2016. PMP for each sample was calculated using the modified enveloping curve. To see the effect of mean and SD the coefficient of variation (Cv) was considered (see Cv = $\frac{SD}{Mean}$ Equation 5).

$$Cv = \frac{SD}{Mean} \qquad \textbf{Equation 5}$$

Therefore, the PMP equation was modified as follows.

$$X_{PMP} = \bar{X}_N (1+ K_m C_v) \qquad \textbf{Equation 6}$$

To analyze the long-term effect of climate change on PMP, percentage differences of each factor were calculated, where P_P = percentage difference of PMP, P_{SD} = percentage difference of SD, P_{Cv} = Percentage difference in coefficient of variation, and P_m = percentage difference in mean. To determine the spatio-temporal changes of these factors, PMP maps were plotted using Arc GIS.

III. RESULTS AND DISCUSSION

A. *Hershfield Statistical Method (S1)*

It should be especially noted that the study is considering only the point PMP values. The enveloping curve for the annual average maximum daily rainfall vs the frequency factor K demonstrates that the Hershfield frequency curve is much higher than the true frequency curve for the Kelani River catchment area. This is also, arguable that the Hershfield K values with low annual average maximum daily rainfall have an extremely high K value when compared with the Hershfield K values having high annual average maximum daily rainfall. This issue was also, a subject of discussion in the Indian context (Sarkar and Maity, 2020c, 2020b). So, this research paper addresses a modified enveloping curve to prevent extremely high values from resulting in overdesigns.

B. *Historical Data Analysis and Modified Hershfield Enveloping Curve (S2)*

With the availability of the historical results, it was identified that only 2 stations recorded a higher rainfall in 1940 than the maximum value of the considered data set in the Kelani River catchment. Therefore, the true K value of those two stations varies as per the grey dots in Figure 1. In the context of the Kelani River catchment, it was recognized that the maximum true K value of 8.85 was gleaned at the average annual maximum daily rainfall of 121.7 mm.

The test provides robust evidence that the maximum value does not act as an outlier in every instance which was also proven in previous studies (Fernando and Wickramasuriya, 2021). Instead, there may be no outliers, maybe one outlier or maybe multiple outliers.

But when observing the historical records of Sri Lanka, it was observed that the K values for the Sri Lankan context have been recognized as 14.8 and 12.03 which is much higher giving the sense that there is a possibility of such an extreme rainfall even in the Kelani River catchment in future. Thus, it is controversial when modifying the Hershfield curve because if one such rainfall occurs the designs may be susceptible to high risk due to underestimation of PMP. To prevent catastrophic consequences, it is vital to contemplate the behaviour of the enveloping curve. It was noticeable that Nedunkeni (12.03) and Balangoda (14.8) have very high true K values compared to other stations in the historical study. Thus, modifying the curve to the highest point (Balangoda) will be reasonable, because this is allowing much space to occur for a very high rainfall from the true K values. Therefore, the Hershfield Curve was modified as a linear horizontal curve as the blue line in Figure 1 with a value Km =14.8 at the mean $\bar{X}$ m = 113.7mm. All the K values having average annual maximum daily rainfall less than $\bar{X}$ m would have a modified value which K = Km= 14.8. Since the $\bar{X}$ m = 113.7mm and only two stations had having annual average maximum daily rainfall less than that, the PMP was modified only to those stations. The results obtained, substantiate that PMP after modifying the curve will reduce the effect of overestimating the PMP at lower annual average maximum daily rainfall. The PMP estimates in the Indian context have also, recorded similar findings (Sarkar and Maity, 2020b, 2020c).

C. *Modified Enveloping Curve for Sri Lankan Context (S3)*

In this section, it is focused on suggesting a modified enveloping curve for the Sri Lankan context. The entire curve was modified to high and low values of annual average maximum daily rainfall. It was discerned that the Hershfield curve is still in a higher position even in higher annual average maximum daily rainfalls. Therefore, it was modified as per the blue line which is a combination of the horizontal and exponential curves in

The test provides robust evidence that the maximum value does not act as an outlier in every instance which was also proven in previous studies (Fernando and Wickramasuriya, 2021). Instead, there may be no outliers, maybe one outlier or maybe multiple outliers.. The reference point used to modify the curve was the values of Nedunkeni for high annual average maximum daily rainfall. PMP values for the Modified curve were much lower than the Hershfield PMP resulting in calculations for the designs in a more economical manner.

D. Detection of Outliers (S4)

Results of the IQR test implied that 6 stations have only one outlier and the PMP will be the same as the Hershfield PMP. However, 2 stations were identified as the presence of two outliers. Thus, the PMP for those two stations will be less than the actual PMP while, the rest has zero outliers representing a high value of PMP than the Hershfield PMP. Therefore, it is judicious to obtain the maximum of the data set as an outlier which will give a false sense of conclusion for the design purposes.

The test provides robust evidence that the maximum value does not act as an outlier in every instance which was also proven in previous studies (Fernando and Wickramasuriya, 2021). Instead, there may be no outliers, maybe one outlier or maybe multiple outliers.

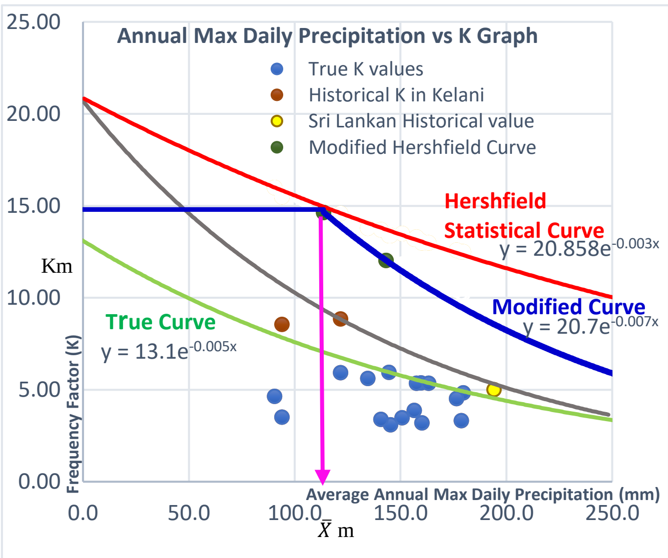

Figure 1. Annual Maximum Daily Rainfall vs K Graph
Source: Author Generated

E. Split sampling and Climate change (S5)

The split sample from 1960-1987 is considered as the dataset 1 (D1) and the sample from 1987-2016 was considered as the dataset 2 (D2) under S5. Results of the split sampling for the change of Mean, SD, Frequency factor, Coefficient of variation (Cv) and PMP were calculated. The results and the interpretation from the split sampling portrayed that the PMP in the later part of 1988-2016 (D2) showed a significant increase in 1960-1987(D1) except for 3 stations (Campion estate, Castlereigh and Bopaththalawa). Stations like Colombo and Norton depict a significant increase in PMP in the later part.

The variation of PMP may be due to several factors. The $X_m = \bar{X}_N + (K \times S_N)$ Equation 1 of S1 depicts the relation of PMP with the Mean, SD and Frequency factors. To see the exact effect, percentage differences for Mean, SD and PMP values were used.

The factors namely, mean and SD have a linear relationship with PMP and the increase of one component of this may influence the increase of PMP. Despite that relationship, it is obvious that, with the increase of mean, the K value decreases. Therefore to, see the effect of SD and mean, the coefficient of variation was used as per $Cv = \frac{SD}{Mean}$ Equation 5 and $X_{PMP} = \bar{X}_N (1+ K_m Cv)$ Equation 6.

The assessment of the variations enumerates that, out of 16 stations, 13 stations are susceptible to an increase in PMP in D2. To find the increase in PMP, a station with high PMP (Colombo) was used. It was examined that $P_m = 5.81\%$, $P_{SD} = 110.97\%$, $P_{Cv} = 99.39\%$, $P_p = 58.49\%$. There is not much variation in the mean for the station Colombo. However, the SD shows a huge increment on D2 compared with D1 showing high P_{SD}. Since it has a higher increment with SD, the increment of Cv will also be higher with a value of PCv = 99.39%. Thus, the P_P was increased to a value of 58.49%, although there is not much difference in the frequency factor. Thus, it is clear that the impact of the SD on the increment of PMP is high. For further analysis to check the effect of the mean, a station with a higher increment in mean Maussakale was considered. $P_m = 36.86\%$, $P_{SD} = 50.53\%$, $P_{Cv} = 10.13\%$, $P_P = 16.9\%$. Since the mean is high the P_{Cv} is low, and also, although P_m is high the P_P is low. It portrays that although there is a high increment of mean, that will not have much effect on the increase of PMP. Thus, it is vital to emphasize that change in SD is the governing factor which directly affects the change of PMP.

The percentage difference of PMP with SD is plotted in Figure 2. It renders that there is a linear relationship between the two factors providing adequate evidence to prove that an increment in SD directly affects the increment in PMP.

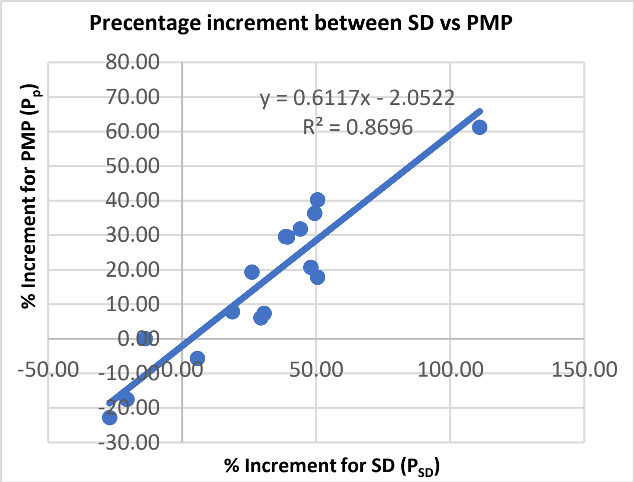

Figure 1. Graph of percentage difference for SD vs PMP
Source: Author Generated

The relation of the PMP and SD will be with the equation, Y= 0.61X - 2.1, where the Y axis is the percentage increment of PMP and the X axis with percentage increment of SD. This would be crucial to determine the long-term effect of change in precipitation on the PMP. If the increment of SD is known in a particular data sample, from the above equation it is possible to determine the change in PMP.

Figure *2* depicts the variation of the percentage differences of Mean, SD, Cv, and PMP. The map of P_P clearly emphasizes that the PMP in the lower catchment area shows a slight increment in PMP in the later part. But considering SD, it could be strongly suggested that, the increment in the lower catchment is huge when comparing D1 and D2. Due to the increment in SD for the lower catchment area, the Cv factor renders similar results, because the increment in the mean is higher for the upper catchment than in the lower catchment. It is vital to emphasize that the SD has a high increment while the PMP also, shows a significant increase in the lower catchment of the Kelani River, concluding that the impact of climate change has not much affected the upper catchment but has affected a considerable amount in the lower catchment this region. However further studies should be conducted in future to examine the impact of climate change by analyzing the percentage increment of PMP.

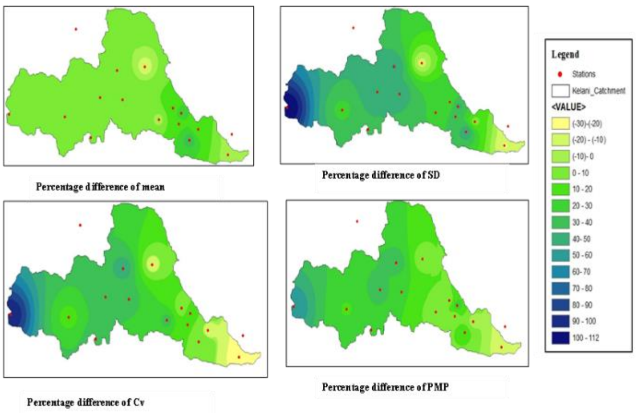

Figure 2. Spatio-temporal variation of Percentage Difference
Source: Author Generated

IV. CONCLUSION AND RECOMMENDATIONS

The original concept of Probable Maximum Precipitation does not incorporate climate change, presently with the global warming impact and with the increment of precipitation, many researchers encapsulate the climate change for PMP.

The S1 was used to determine the Hershfield enveloping curve concluding the need to modify the enveloping curve as in S2. This was the main objective of this study. This is used to estimate the basin-specific 24-hour PMP values for the catchment. The historical values were taken into account in two aspects, mainly the Historical values in the aspect of the Kelani River catchment and with the Sri Lankan context. K values in high Annual Average Maximum Daily Precipitation were modified using the Sri Lankan context in S3. It is noticeable that the curve was below the Hershfield curve for high Annual Average Maximum Daily Precipitations. Thus, it is suggested that it would be the most suitable enveloping curve for the Sri Lankan context.

S4, emphasizes the impact of outliers on the PMP. After analyzing the results, it is vital to consider the effect of the outlier into account rather than considering the maximum of each station as an outlier as per the Hershfield principle.

The data sample of 57 years was split into two datasets D1 and D2 under S5, to analyze the long-term effect of PMP. Mean, SD, Cv and PMP for each split sample were determined for D1 and D2. It was observed that PMP for D2 was higher than D1 for 13 stations out of 16. A further study was conducted to find the factors which affect the increment of PMP. PMP depends on mean and SD. Seeking for the factor it was observed that SD has increased compared with D1 although the difference in mean was negligible. The impact was determined using the percentage differences. In conclusion, although the SD differs by a considerable amount the change of PMP is very low when compared with the change in SD. Therefore, it is vital to consider the change in SD when considering the long-term effect of PMP change. The plot of the P_P and SD deviation provides a linear relationship of providing robust, evidence that the change of PMP directly incorporates with the change of SD. Thus, it is concluded that from the relationship, although there is a very high increase in SD, the difference in PMP is small.

The Percentage difference maps emphasize the variation of the SD along with PMP in a spatial-temporal aspect. The lower catchment has a higher increase in SD resulting in a slight increase in PMP. The upper catchment shows an increment in percentage difference for the mean. Concluding the arguments, it is vital to note that although the SD has a significant increase, the variation of PMP is much less.

REFERENCES

Bank, A.D. (2021) 'Climate risk country profile'.

Carter, T.R. (1996) 'Assessing Climate Change Adaptations : The IPCC Guidelines'.

Fernando, W.C.D.K. and Wickramasuriya, S.S. (2021) 'Concept of Threshold in the Estimation of Probable Maximum Precipitation: Hershfield's Method Revisited', *Journal of Hydrologic Engineering*, 26(3), pp. 1–8. Available at: https://doi.org/10.1061/(asce)he.1943-5584.0002045.

Ministry of Mahaweli Development and Environment (2016) *National Adapotation plans for Climate Change Impacts in Sri Lanka*. Available at: https://www4.unfccc.int/sites/NAPC/Documents NAP/National Reports/National Adaptation Plan of Sri Lanka.pdf.

Sarkar, S. and Maity, R. (2020a) 'Estimation of Probable Maximum Precipitation in the context of climate change', *MethodsX*, 7. Available at: https://doi.org/10.1016/j.mex.2020.100904.

Sarkar, S. and Maity, R. (2020c) 'Increase in probable maximum precipitation in a changing climate over India', *Journal of Hydrology*, 585(March), p. 124806. Available at: https://doi.org/10.1016/j.jhydrol.2020.124806.

Ratnayake N. (Ed.), Proceedings of the Civil Engineering Research for Industry Symposium 2011 (pp. 73-78).

Thanh Thuy, L.T., Kawagoe, S. and Sarukkalige, R. (2019) 'Estimation of probable maximum precipitation at three provinces in Northeast Vietnam using historical data and future climate change scenarios', *Journal of Hydrology: Regional Studies*, 23(August 2018), p. 100599. Available at: https://doi.org/10.1016/j.ejrh.2019.100599.

Walfish, S. (2006) 'A review of statistical outlier methods', *Pharmaceutical Technology*, 30(11), pp. 82–86.

Wickramasuriya, S.S. and Fernando, W.C.D.K. (2012) 'Challenges in Dam Safety and Extreme Rainfall Estimation in Relation to Sri Lanka', XXXXV(01).

World Meteorological Organization (2009) *22. Manual for Estimation of Probable Maximum Precipitation, Operational Hydrology Report 1, 2nd edition, Publication 332*.

ACKNOWLEDGMENT

The authors wish to thank the officials of the Department of Civil Engineering, General Sir John Kotelawala Defence University for providing the opportunity to conduct this research.

AUTHOR BIOGRAPHIES

KMSS Gunathilaka has completed a BSc (Hons) in Civil Engineering from General Sir John Kotelawala Defence University, Sri Lanka. Her area of interest in research is Hydrology.

Prof. WCDK Fernando is a Professor attached to the Department of Civil Engineering, General Sir John Kotelawala Defence University. Her primary research field is Engineering Hydrology with special reference to extreme rainfall analysis.

Prof SS Wickramasuriya is a Professor attached to the Department of Civil Engineering, General Sir John Kotelawala Defence University. His Primary research field is Dam safety, Hydrological extremes and stochastic Hydrology.