Evaluation of Alternative Approach for Missing Rainfall Data Filling in Kuantan River Basin

Shirline Leong Ai Lin[1, a)], Jaqueline Isabella Anak Gisen[1, b)], Muhammad Amiruddin Zulkifli[1,c)], Muhamad Amin Kamaruddin[1,d)] and Shairul Rohaziawati Samat[1,e)]

[1]Faculty of Civil Engineering Technology, Universiti Malaysia Pahang, Persiaran Tun Khalil Yaakob, 26300 Kuantan, Malaysia

[a)] *Corresponding author: shirlineleong95@outlook.com*
[b)] *isabella@ump.edu.my*
[c)] *amiruddinzulkifli7@gmail.com*
[d)]*aminkamaruddin@gmail.com*
[e)] *shairul@ump.edu.my*

Abstract. Water authority has been dealing with missing precipitation data of Kuantan River Basin (KRB) for decades and with recent flash flood events in the state of Pahang has highlighted the importance of climate data in flood prediction to reduce the severity of flood damage in future. However, climate data collected on site often contains gaps affecting the quality of rainfall data resulting in inaccuracy in analysis results. The main objective of this study is to evaluate the relationship between observed and remotely sensed (TRMM) rainfall data whether there is any established relationship that is fit to act as an alternative approach to fill missing observed rainfall data for KRB. The rainfall data were collected from three different sources which are observed rainfall data from Department of Irrigation and Drainage (DID), meteorological rainfall data from Malaysia Meteorological Department (MMD) and TRMM rainfall data from NASA website. It was found that correlation between the untreated observed rainfall data and remotely sensed rainfall data is not strong enough to be an alternative approach. The most noticeable finding was from Kg Sg Soi station where the correlation coefficient between TRMM and DID observed rainfall was found to be 0.56 and the relationship between TRMM and MMD rainfall data appear to have better correlation with a coefficient of 0.57. However, when rainfall data was analysed by month, correlation was as high as 0.74 which proved that correlation is easier to be established in months during wet season. Subsequently, XLSTATS Software was used to input the missing *observed* rainfall values for 8 active rainfall stations to find out the best imputation method for KRB's missing observed rainfall data. To assess the method's performance, the results were compared to the conventional approach which is station average method. The outcomes for this study have proved that Replace by Mean, MCMC and Nearest Neighbor method are the best approach to estimate the missing rainfall data for all the station in KRB. This study's findings provide a full observed rainfall dataset and the best imputation approaches for all 8 active rainfall stations that can be utilized for future hydrological studies.

Keywords: Missing Rainfall Data, TRMM Satellite Precipitation Product, Multiple Imputation Method.

INTRODUCTION

Access to whole rainfall dataset is vital, particularly in Malaysia that is prone to floods during wet season and drought during dry season. During the inter-monsoon and NE monsoon seasons, floods are very common. Rainfall events lasting many days result in massive volume of stormwater, causing significant floods in several parts of the country [1]. Another common issue in Malaysia is drought which has complicated effects on a variety of economic sectors, including water resources, agriculture, and natural ecosystems [2]. It is a major threat to long-term growth and urban development of the country.