

Application of single imputation method in estimating missing value for ozone data

ABSTRACT

Handling missing value problems is an essential procedure in research. Air pollution data acquired utilising automated instruments frequently consist of missing values, which can lead to biased conclusions and degrade the efficiency of the data. The application of the imputation method plays a crucial role in handling this problem. In this study, three single imputation methods were considered, where a simulation study was performed to compare the three methods of imputation, namely linear interpolation, cubic interpolation, and mean-before-after (MBA) method. The performance of each method was measured using two selected performance indicators to obtain the best method that can be used to handle missing values in air pollution data. The MBA method was found to be superior compared to the other two methods. The MBA method was then utilised on the ozone data from the industrial area of Petaling Jaya, Malaysia, which has missing values. Weibull distribution was considered to describe the ozone data. The parameter estimates of ozone data imputed with the MBA imputation method together with its standard deviation indicated better estimation as the values obtained were smaller compared to the estimates of ozone data with missing values. Therefore, the findings of this study showed that, via simulation study, the most effective method for handling missing values was the MBA method.