

Factor Analysis Affected DBD Cases In Central Java Province Using Random Forest Regression and Spatial Autoregressive Models (SAR)

Nur Adila¹

¹14611031@students.uii.ac.id

Student at Departement of Statistics

Islamic University of Indonesia | Kaliurang Street KM 14,5 Yogyakarta 55584

ABSTRACT

The territory of Indonesia began in 1968 until 2009, according to the World Health Organization (WHO) recorded as the country with the highest dengue cases in Southeast Asia and is one of the deadliest diseases. One of the provinces included in the 10 provinces with the highest DHF level in Indonesia is Central Java Province. The large number of dengue cases in Central Java is caused by several factors including population density, number of health workers, number of health facilities, number of PHBS families, area height and amount of rainfall. This study aims to determine the factors that significantly influence the dengue cases in Central Java Province in 2016 by using the Random Forest Regression method and knowing the prediction models produced using SAR. The purpose of using the random forest regression method is to standardize, so that the MSE results obtained are smaller. Based on the results of random forest regression analysis before standardization or after standardization, the number of trees selected is 200 trees with the m being tested is 2.44 and 1.22. variables that influence DHF cases before standardization, namely the number of PHBS families. While the variables that influence DHF cases after standardization are population density. In this study used Spatial Autoregressive (SAR) analysis to find out the results of the prediction model. This is because the DHF case is very likely to be affected by the spatial or position of other areas that are around it. Results of SAR models that can be formed

are $\hat{y}_i = 295,185 + 0,44 \sum_j^n W_{ij} Y_j - 0,0708 \text{Kepadatan Penduduk} + 0,2751 \text{Jumlah Tenaga kesehatan} - 0,1443 \text{Curah} + \varepsilon$

Keywords : *case of dengue hemorrhagic fever, Radom Forest Regression, SAR.*