

Effect of monsoonal clustering for pm10 concentration Prediction in Keningau, Sabah using principal component analysis

ABSTRACT

Particulate matter (PM) has caught scientific attention in scientific research due to its harmful effect on human health. While prediction is essential for future development in Keningau, temporal clustering in Keningau has yet to be studied. Thus, this research aims to determine whether monsoonal clustering is required for meteorological and pollutant concentration data collected in Keningau. Missing data is first imputed using Nearest Neighbour Method (NNM). Then, wind direction and wind speed are converted into northern (Wy) and eastern (Wx) component of wind speed. Data is then temporal clustered based on monsoonal season (NEM, IM4, SWM, IM10). Both clustered and unclustered data are analysed using principal component (PC) analysis (PCA). The findings revealed that humidity in PC1 with average EV (explained variation) of 93.92 ± 0.52 contribute the most variation of PM10, followed by Wx in PC2 with average EV of 3.51 ± 0.48 . Regression analysis shows that humidity and PM10 are negatively moderate to strongly correlated except for IM4 (intermonsoon April), which may be due to dry climate during the season. As for Wx, it has weak correlation with PM10. This may be due to location of Keningau at western part of Crocker range. However, the spread of PM10 due to eastern wind causes weak to zero correlation. Due to consideration of dry climate as revealed by the findings from IM4 cluster, there is need for data collected by Keningau to be clustered by monsoon.