

## **Assessment of various probability distributions to model Pm 10 concentration for industrialized area in peninsula Malaysia: A case study in shah alam and nilai**

### **Abstract**

In Malaysia, the predominant air pollutants are suspended particulate matter (SPM) and nitrogen dioxide (NO<sub>2</sub>). This research focused on PM<sub>10</sub> as they may trigger harm to human health as well as environment. Six distributions, namely Weibull, log-normal, gamma, Rayleigh, Gumbel and Frechet were chosen to model the PM<sub>10</sub> observations at two industrial areas: Nilai and Shah Alam. One-year period hourly average data for 2006 and 2007 were used in this research. For parameters estimation, method of maximum likelihood estimation (MLE) was selected. Four performance indicators that are mean absolute error (MAE), root mean squared error (RMSE), coefficient of determination (R<sup>2</sup>) and prediction accuracy (PA), were applied to determine the goodness-of-fit criteria of the distributions. The best distribution that fits the PM<sub>10</sub> observations was found to be gamma distribution for Nilai whereas for Shah Alam, log-normal distribution is more appropriate. The probabilities of the exceedences concentration were calculated and the return period for the coming year was predicted from the cumulative density function (cdf) obtained from the best-fit distributions. For the 2006 data, Nilai was predicted to exceed 150 µg/m<sup>3</sup> for 2.7 days in 2007 with a return period of one occurrence per 137 days. Shah Alam was predicted to exceed 150 µg/m<sup>3</sup> for 5.9 days in 2007 with a return period of one occurrence per 62 days. Both areas do not exceed the MAAQG of 150 µg/m<sup>3</sup> based on 2007 data.

### **Keywords**

Exceedences; Particulate matter; Performance indicators; Probability distributions; Return period; Statistical analysis