Towards on-site implementation of multi-step air pollutant index prediction in Malaysia industrial area: Comparing the NARX neural network and support vector regression

ABSTRACT

Malaysia has experienced public health issues and economic losses due to air pollution problems. As the air pollution problem keeps increasing over time, studies on air quality prediction are also advancing. The air quality prediction can help reduce air pollution's damaging impact on public health and economic activities. This study develops and evaluates the Nonlinear Autoregressive Exogenous (NARX) Neural Network and Support Vector Regression (SVR) for multi-step Malaysia's Air Pollutant Index (API) prediction, focusing on the industrial areas. The performance of NARX and SVR was evaluated on four crucial aspects of on-site implementation: Input pre-processing, parameter selection, practical predictability limit, and robustness. Results show that both predictors exhibit almost comparable performance, in which the SVR slightly outperforms the NARX. The RMSE and R2 values for the SVR are 0.71 and 0.99 in one-step-ahead prediction, gradually changing to 6.43 and 0.68 in 24-step-ahead prediction. Both predictors can also perform multi-step prediction by using the actual (non-normalized) data, hence are simpler to be implemented on-site. Removing several insignificant parameters did not affect the prediction performance, indicating that a uniform model can be used at all air quality monitoring stations in Malaysia's industrial areas. Nevertheless, SVR shows more resilience towards outliers and is also stable. Based on the trends exhibited by the Malaysia API data, a yearly update is sufficient for SVR due to its strength and stability. In conclusion, this study proposes that the SVR predictor could be implemented at air quality monitoring stations to provide API prediction information at least nine steps in advance.