

MODELING MULTIVARIABLE AIR POLLUTION DATA IN MALAYSIA USING VECTOR AUTOREGRESSIVE MODEL

(Memodelkan Data Pencemaran Udara Multipemboleh Ubah di Malaysia Menggunakan Model Autoregresi Vektor)

ULYA BT ABDUL RAHIM & NURULKAMAL MASSERAN*

ABSTRACT

In this study, the vector autoregressive (VAR) model was used to model and forecast the multivariable air pollution data in Klang area. Stationary test, Hannan–Quinn evaluation criteria, Granger causality test, R^2 coefficient and Root Square Mean Error (RMSE) measurements have been conducted to get the best model and will be used in forecasting. The VAR (7) model is found to be the best model with the highest R^2 and lowest RMSE value recorded for each dependent pollutant variable. Based on the fitted VAR (7) model, the VAR model is able to describe the dynamic behavior of multivariable air pollution data of Klang. Forecasts of up to 12 days ahead were constructed with confidence intervals. The VAR model found to provides good forecast accuracy on the data.

Keywords: air-pollution modeling; VAR model; forecasting

ABSTRAK

Dalam kajian ini, model autoregresi vektor (VAR) digunakan untuk memodel dan meramal data pencemaran udara multipemboleh ubah di kawasan Klang. Ujian kepegungan, kriterium penilaian Hannan–Quinn, ujian kebersebaban Granger, pekali R^2 dan ukuran ralat min kuasa dua (RMSE) telah dijalankan untuk mendapatkan model terbaik dan akan digunakan bagi tujuan peramalan. Model VAR (7) dikenal pasti sebagai model terbaik dengan nilai pekali R^2 tertinggi dan nilai RMSE yang terendah untuk setiap pemboleh ubah pencemar bersandar. Berdasarkan model VAR (7) yang disuaikan, model VAR didapati mampu untuk memerihalkan tingkah laku dinamik data pencemaran udara multipemboleh ubah di Klang. Ramalan sehingga 12 hari ke depan telah dijalankan beserta maklumat selang keyakinan bagi model VAR(7). Model VAR didapati boleh memberikan ketepatan ramalan yang baik terhadap data.

Kata kunci: pemodelan pencemaran udara; model VAR; peramalan

References

- Al-Dhurafi N.A., Masseran N. & Zamzuri Z.H. 2018. Compositional time series analysis for Air Pollution Index data. *Stochastic Environmental Research and Risk Assessment* **32**(10): 2903-2911.
- Ary Pani D. 2016. Pemodelan Pencemaran Udara Menggunakan Metode Vector Autoregressive (VAR) di Provinsi Riau. *Jurnal Sains dan Teknologi Industri* **13**(2): 160-167.
- Bai L., Wang J., Ma X. & Lu H. 2018. Air pollution forecasts: An overview. *International Journal of Environmental Research and Public Health* **15**(4): 1-44.
- Barnston A.G. 1992. Correspondence among the Correlation, RMSE, and Heidke Forecast Verification Measures; Refinement of the Heidke Score. *Weather and Forecasting* **7**: 699-709.
- Brooks C. 2019. *Introductory econometrics for finance*. Cambridge: Cambridge University Press
- Cameron A.C. & Windmeijer F.A. 1997. An R-squared measure of goodness of fit for some common nonlinear regression models. *Journal of Econometrics* **77**(2): 329-342.
- Gass K., Klein M., Sarnat S., Winquist A., Darrow L., Flanders W., Chang H., Mulholland J., Tolbert P. & Strickland M. 2015. Associations between ambient air pollutant mixtures and pediatric asthma emergency department visits in three cities: a classification and regression tree approach. *Environmental Health* **58**:1–14
- Liew V.K.S. 2004. Which lag length selection criteria should we employ? *Economics Bulletin* **3**(33): 1-9.

- Masseran N. 2017. Modeling fluctuation of PM₁₀ data with existence of volatility effect. *Environmental Engineering Science* **34**(11): 816-827
- Masseran N. & Razali A.M. 2016. Modeling the wind direction behaviors during the monsoon seasons in Peninsular Malaysia. *Renewable and Sustainable Energy Reviews* **56**:1419–1430
- Masseran N., Razali A.M., Ibrahim K. & Latif M. T. 2016. Modeling air quality in main cities of Peninsular Malaysia by using a generalized Pareto model. *Environmental Monitoring and Assessment* **188**(1), Article number 65: 1-12.
- Masseran N., Razali A.M., Ibrahim K. & Wan Zin W. Z. 2012. Evaluating the wind speed persistence for several wind stations in Peninsular Malaysia. *Energy* **37**(1): 649-656.
- Masseran N., Razali A.M., Ibrahim K., Zaharim A., & Sopian K. 2013. Application of the single imputation method to estimate missing wind speed data in Malaysia. *Research Journal of Applied Sciences, Engineering and Technology* **6**(10): 1780-1784.
- Omasa K. 2002. Diagnosis of stomatal response and gas exchange of trees by thermal remote sensing. *Tokyo: Springer*.
- Pfaff B. 2008. VAR, SVAR and SVEC models: Implementation within R package vars. *Journal of Statistical Software*, 27(4), 1-32.
- World Health Organization. 2006. WHO Air quality guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide: global update 2005: summary of risk assessment. Geneva: World Health Organization.

Statistics Programme
Faculty of Science and Technology
Universiti Kebangsaan Malaysia
43600 UKM Bangi
Selangor DE, MALAYSIA
*E-mail: p89235@siswa.ukm.edu.my, kamalmsn@ukm.my**